Spring 5-10-2014

Molecular Dynamics Simulations Towards The Understanding of the Cis-Trans Isomerization of Proline As A Conformational Switch For The Regulation of Biological Processes

Hector Velazquez

MOLECULAR DYNAMICS SIMULATIONS TOWARDS THE UNDERSTANDING OF THE CIS-TRANS ISOMERIZATION OF PROLINE AS CONFRONTATIONAL SWITCH FOR THE REGULATION OF BIOLOGICAL PROCESSES

by

HECTOR VELAZQUEZ

Under the Direction of Dr. Donald Hamelberg

ABSTRACT

Pin1 is an enzyme central to cell signaling pathways because it catalyzes the *cis–trans* isomerization of the peptide ω-bond in phosphorylated serine/threonine-proline motifs in many proteins. This regulatory function makes Pin1 a drug target in the treatment of various diseases. The effects of phosphorylation on Pin1 substrates and the basis for Pin1 recognition are not well understood. The conformational consequences of phosphorylation on Pin1 substrate analogues and the mechanism of recognition by the catalytic domain of Pin1 were determined using molecular dynamics simulations. Phosphorylation perturbs the backbone conformational space of Pin1 substrate analogues. It is also shown that Pin1 recognizes specific conformations of its substrate by conformational selection. Dynamical correlated motions in the free Pin1 enzyme are present in the enzyme of the enzyme–substrate complex when the substrate is in the transition

state configuration. This suggests that these motions play a significant role during catalysis. These results provide a detailed mechanistic understanding of Pin1 substrate recognition that can be exploited for drug design purposes and further our understanding of the subtleties of post-translational phosphorylation and *cis–trans* isomerization.

Results from accelerated molecular dynamics simulations indicate that catalysis occurs along a restricted path of the backbone configuration of the substrate, selecting specific subpopulations of the conformational space of the substrate in the active site of Pin1. The simulations show that the enzyme–substrate interactions are coupled to the state of the prolyl peptide bond during catalysis. The transition-state configuration of the substrate binds better than the cis and trans states to the catalytic domain of Pin1. This suggests that Pin1 catalyzes its substrate by noncovalently stabilizing the transition state. These results suggest an atomistic detail understanding of the catalytic mechanism of Pin1 that is necessary for the design of novel inhibitors and the treatment of several diseases. Additionally, a set of constant force biased molecular dynamics simulations are presented to explore the kinetic properties of a Pin1 substrate and its unphosphorylated analogue. The simulations indicate that the phosphorylated Pin1 substrate isomerizes slower than the unphosphorylated analogue. This is due to the lower diffusion constant for the phosphorylated Pin1 substrate.

INDEX WORDS: Pin1, Accelerated molecular dynamics, Cis-trans isomerization, PPIases, Enzyme dynamics, Peptidyl-prolyl *cis-trans* isomerase, Molecular switches, Protein regulation

MOLECULAR DYNAMICS SIMULATIONS TOWARDS THE UNDERSTANDING OF THE CIS-TRANS ISOMERIZATION OF PROLINE AS CONFRONTATIONAL SWITCH FOR THE REGULATION OF BIOLOGICAL PROCESSES

by

HECTOR VELAZQUEZ

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the College of Arts and Sciences

Georgia State University

2014

MOLECULAR DYNAMICS SIMULATIONS TOWARDS THE UNDERSTANDING OF THE CIS-TRANS ISOMERIZATION OF PROLINE AS CONFRONTATIONAL SWITCH FOR THE REGULATION OF BIOLOGICAL PROCESSES

by

HECTOR VELAZQUEZ

Committee Chair: Donald Hamelberg

Committee: Ivaylo Ivanov

Aimin Liu

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

May 2014

DEDICATION

This dissertation is dedicated to my wife and parents

ACKNOWLEDGEMENTS

Although many people have supported me throughout my doctoral research, I could not have accomplished the work presented in this dissertation without the guidance provided by my advisor and mentor Dr. Donald Hamelberg. I am eternally grateful to Dr. Hamelberg for giving me the opportunity to do my dissertation research under his guidance. Dr. Hamelberg is an outstanding professor and advisor from whom I have learned much about academia and research.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **v**

LIST OF TABLES ..... **x**

LIST OF FIGURES ..... **xi**

1 INTRODUCTION **1**

1.1 References **4**

2 CONFORMATIONAL SELECTION IN THE RECOGNITION OF PHOSPHORYLATED SUBSTRATES BY THE CATALYTIC DOMAIN OF PIN1 ..... **7**

2.1 Abstract **7**

2.2 Introduction **8**

2.2 Methods **11**

2.2.1 Accelerated Molecular Dynamics of Pin1 substrate analogue ***11***

2.2.2 Molecular Dynamics Simulations of Pin1 and Pin1 substrate complexes . ***12***

2.3 Results and Discussion **13**

2.3.1 Phosphorylation effects on the equilibriumof Pin1 substrate analogues ... ***13***

2.3.2 Conformational Selection in Pin1 Recognition ***16***

2.3.3 Dynamical and Correlated Motions in Pin1 ..... ***21***

2.4 Concluding Remarks **23**

2.5 Supporting Information **23**

2.6 Acknowledgments 24

2.7 References 24

2.8 Figures 32

3 CONFORMATION DIRECTED CATALYSIS AND COUPLED ENZYME-SUBSTRATE DYNAMICS IN PIN1 PHOSPHORLATION DEPENDENT CIS-TRANS ISOMERASE 36

3.1 Abstract 36

3.2 Introduction 37

3.3 Computational Methods 39

3.4 Results and Discussion 41

3.4.1 Conformation Directed Catalysis by Pin1 41

3.4.2 Interactions between Pin1 and the phosphate moiety are key to stabilizing the transition state 43

3.4.3 Several intermolecular Pin1-substrate and intramolecular Pin1 interactions are coupled to the dynamics of the substrate 46

3.4.4 Pin1 preferentially binds the transition state configuration of its substrate 48

3.5 Conclusions 49

3.6 Acknowledgement 50

3.7 References 50

3.8 Figures 57

4 KINETIC RATES AND DYNAMICAL EFFECTS OF PHOSPHORYLATION ON PEPTIDYL PROLYL CIS-TRANS ISOMERIZATION FROM CONSTANT FORCE MOLECULAR DYNAMICS 60

4.1 Abstract 60

4.2 Introduction 61

4.3 Computational Methods 64

4.3.1 Accelerated molecular dynamics of prolyl peptide substrates 64

4.3.2 Simulations of the cis to trans transition under a constant pulling force .. 66

4.3.3 Construction of the Survival Probability Distributions 67

4.3.4 Fitting to the Dudko-Hummer-Szabo (DHS) Model 67

4.3.5 Calculation of the autocorrelation functions at 300K 68

4.4 Results and Discussion 68

4.4.1 The peptidyl-prolyl bond angle is correlated to the Cα-Cα distance 68

4.4.2 Cis to trans isomerization under tension 69

4.4.3 Extracting the survival time at zero force with the Dudko-Hummer-Szabo (DHS) model 70

4.4.4 Extracting the survival time at zero force with the Bell Model 71

4.4.5 Diffusion Accounts for the difference in the kinetic rate constants of the phosphorylated and unphosphorylated polypeptide substrates 71

4.5 Conclusions 72

4.6 References 73

4.7 Figures 80

4.8 Tables 83

5 CONCLUDING REMARKS AND FUTURE STUDIES 84

LIST OF TABLES

Table 1. Data of constant pulling force simulations 83

Table 2. Fitted Parameters for the DHS model 83

LIST OF FIGURES

Figure 1. Pin1 catalyzes cis-trans isomerization of the ω-bond angle of proline when the preceding residue is a phosphorylated serine or phosphorylated threonine. The ω-bond angle is shown in yellow when it is in the trans (left) and cis (right) configuration..... 32

Figure 2. Contour plot in kcal/mol of the conformational space of unphosphorylated serine and phosphorylated serine in the free substrate, Ace-Ala-Ala-Ser/pSer-Pro-Phe-Nme. (A) The Ramachandran plot (top) of Ser and ω-ψ plot (bottom). 33

Figure 3. Contour plot in kcal/mol of the conformational space of Pro in the free phosphorylated and unphosphorylated substrate analogue. (A) The φ-ψ (top) and ω-ψ plot (bottom) of proline and the preceding ω-bond angle of the pSer-Pro motif..... 33

Figure 4. Structure and binding site of the catalytic domain of Pin1. (A) X-ray crystal structure (PDB ID 2Q5A) of the catalytic domain of Pin1 (shown as white), with a peptidomimetic in the active site. The active residues are shown using surface representation. . (B) The residues that make up the active sites. (C) The peptidomimetic inhibitor is shown interacting with the proline binding pocket and the phosphate binding residues (Lys 63, Arg 68, and Arg 69), shown as blue. (D) The substrate analogue, Ace-Ala-Ala-pSer-Pro-Phe-Nme, is shown in the active site interacting with the proline binding pocket and the phosphate binding residues(Lys 63, Arg 68, and Arg 69), shown as blue. 33

Figure 5. Contour plot in kcal/mol of the φ-ψ space or Ramachandran plot of the pSer-Pro motif of the substrate analogue, Ace-Ala-Ala-pSer-Pro-Phe-Nme, in the active site of the catalytic domain of Pin1 for pSer (top) and Pro (below), (A) when the ω-bond angle of the pSer-Pro motif is in the trans configuration and (B) when the ω-bond angle of the pSer-Pro motif is in the cis configuration..... 34

Figure 6. Contour plot in kcal/mol of the ω-ψ space of (A) pSer and (B) Pro of the pSer-Pro motif of the substrate analogue, Ace-Ala-Ala-pSer-Pro-Phe-Nme, when bound in the active site of the catalytic domain of Pin1..... 34

Figure 7. (A) Probability distribution of the first principal component (PC1) of the Pin1-substrate complex when the ω-bond angle of the pSer-Pro motif of the substrate is in the trans (blue), cis (green), and transition states (red) configurations and free Pin1 35

Figure 8. Dynamical cross correlated motions of Pin1 residues in the (A) free Pin1, and when bound to the substrate in the (B) trans, (C), cis, and (D) transition state configurations. The main correlated motions of Pin1 are labeled as m1, m2, and m3, and the re residues and domains involved in these motions are depicted on the right. 35

Figure 9. Pin1 and the substrate analogue (AApSPF) in the active site. A close-up of the active site is shown and the relevant backbone dihedral angles of the substrate are also defined. The phosphate-binding site is colored using green transparent surface representation and the residues are shown using the green stick representation. The residues shown in green are Lys 63, Arg 68, Arg 69, and Ser 154. The rest of the active site is shown using a white transparent surface representation and stick model. 57

Figure 10. The phase space of the -pSer-Pro- motif in the substrate analogue free in solution and in the active site of Pin1. Free energy landscapes in kcal/mol along ω-ψ of Pro (A) and pSer (B) free in solution and Pro (C) and pSer (D) in the active site of Pin1. Free energy landscapes in kcal/mol along ω-ψ of Pro (A) and pSer (B) free in solution and Pro (C) and pSer (D) in the active site of Pin1. The dashed lines with arrows in (C) and (D) depict the lowest free energy paths along the backbone of Pro and pSer during catalysis. 58

Figure 11. Free energy of cis-trans isomerization of the substrate in solution and the active site of Pin1. (A) Free energy profiles in kcal/mol along the ω-bond angle of the –pSer-Pro- motif in the substrate analogue for the uncatalyzed isomerization reaction (b lack) and the catalyzed isomerization reaction by Pin1(red). (B) Thermodynamic cycle connecting the isomerization process in solution and active site of Pin1 with the free energies of binding the cis (c), transition (ts) and trans (t) states of the substrate. 58

Figure 12. Dynamical coupling between the phosphate group of the substrate and Pin1 phosphate-binding site during catalysis. The residues in the phosphate binding site of Pin1 and the interactions made with the substrate (A). 2D free energy profiles in kcal/mol along the ω-bond of the -pSer-Pro- motif of the substrate and the distance between the guanidinium carbon of Arg 68 and the phosphorus atom of the phosphate group (B), the distance between the guanidinium carbon of Arg 69 and the phosphorus atom of the phosphate group (C), and the distance between the side chain nitrogen atom of Lys 63 and the phosphorus atom of the phosphate group (C)..... 59

Figure 13. Dynamical coupling between the phosphate group of the substrate and Ser154, in the active site of Pin1, during catalysis. 2D free energy profiles in kcal/mol along the ω-bond of the -pSer-Pro- motif of the substrate and the distance between the side ch oxygen atom of Ser 154 and the phosphorus atom of the phosphate group (A). Ser 154 in the active site of Pin1 and the substrate analogue in the active site are shown (B). 59

Figure 14. Dynamical coupling between the substrate and Pin1 during catalysis. 2D free energy profiles in kcal/mol along the ω-bond of the -pSer-Pro- motif of the substrate and the distance between the sulfur atom of Cys113 and the carbonyl oxygen atom of the pe ptide ω-bond of the substrate (A), the distance between the sulfur atom of Cys113 and the ε-nitrogen of

His59 (B), the distance between the side chain nitrogen atom of Gln131 and the carbonyl oxygen atom of the peptide ω-bond of the substrate (C), and the distance between the side chain oxygen atom of Gln131 and the side chain oxygen atom of Thr152 (D). 60

Figure 15. Probability distributions of free energies of binding the substrate in different configurations. Free energies in kcal/mol of binding the transitions state (black), cis-α (red) and trans-α (blue) configurations in the active site of Pin1. 60

Figure 16. Prolyl cis-trans isomerization about the omega bond (ω). The alpha carbons of the X residue and the prolyl residue are represented as orange spheres. The biasing force for the constant force biased simulations were applied to the alpha carbons of the X residue and the prolyl residue are represented as orange spheres. The biasing force for the constant force biased simulations were applied to the alpha carbons of the X and prolyl residues as illustrated by the arrows. 80

Figure 17. Free Energy Profiles showing the coupling of the omega bond to the c-alpha - c-alpha distance. Panel A shows how the energy fluctuates with the distance between the alpha carbons of the X and prolyl residues. Panel B shows the coupling between the omeg a bond and the distance between the alpha carbons of the X and prolyl residues. 81

Figure 18. Constant force bias simulation of the X = Ser substrate where no cutoff distance is assigned to end the simulation. The distance being monitored is the distance between the alpha carbons of the X and prolyl residues. The force for this simulation is 60 kcal/mol*Å. 81

Figure 19. Survival Probability distributions for X = Ser (A) and X = pSer (B) substrates. 81

Figure 20. Kinetic fit of the constant force biased simulations to the DHS model (A) and the Bell model (B) 82

Figure 21. Free Energy Profiles of prolyl cis-trans isomerization for the X-Pro motif where X=Ser (black) or pSer (red). 82

Figure 22. Autocorrelation function of unphosphorylated (black) and phosphorylated (red) polypeptide substrates at 300K. 82

1 INTRODUCTION

The publication of Watson and Crick's paper of the structure of DNA marked the beginning of a revolution in chemical research[1]. This achievement paved the way for a full understanding the underlying chemical processes involved in life. Since then, proteins have been established as the "work horses" of the cell, carrying out many of the essential functions associated with living organisms[2]. This realization leads to more questions such as: how are proteins regulated and more importantly why is this necessary? The biochemical literature indicates that proteins actually regulate other proteins via post-translational modifications so that the proper biological functions are executed at the proper time[3]. For example, histones are regulated via acetylation and deacetylation by histone acetyl transferase and histone deacetylase, respectively, as part of gene regulation[4,5]. More specifically, histones are the proteins responsible for packing DNA into nucleosomes[6]. This type of regulation assures that DNA replication is only carried out at the proper times during the cell cycle. Another form of protein regulation is a post-translational modification known as glycosylation. The purpose of this regulation is to ensure that certain proteins fold properly thereby preserving their functions[3, 6]. Many biological molecules are influenced by this type of regulation including enzymes, antifreeze proteins, and hormones [7, 8]. While much is known about protein regulation via post-translational modifications such as glycosylation, much more is still being discovered and reported in the biochemical literature.

One area where much remains to be discovered is the post-translational phosphorylation of proteins[9]. One reason for the lack of knowledge in this area is that phosphorylation is the most common type of post-translational modification indicating that it plays many different roles in biological systems. This makes it it difficult to acquire a global view of how it regulates biological functions. What is truly remarkable about post-translational phosphorylation of

proteins is its specificity. This type of modification only occurs at serine, threonine, or tyrosine residues in proteins. Despite this strict specificity, phosphorylation is an extremely diverse way to regulate proteins. Phosphorylation can accomplish this diversity because of the ways in which it influences protein structure. Since a phosphate group is electrically charged, its presence in a protein radically changes the protein's solvation structure. This has a cascading effect where the hydrogen bonding patterns on the protein's surface and salt bridge interactions are changed. This perturbation of the interaction patterns will trigger changes in the protein's conformational backbone structure[10] which can in turn render proteins active or inactive due to changes in the secondary structure of the protein[11]. The phosphate group is also sensitive to micro pH environments in the cell since it can accommodate multiple protonation states in the pH range of known biological systems thereby adding another level to its versatility in protein regulation[12]. The mere size of a phosphate group could render the protein it modifies incapable of interacting with the binding pocket of certain proteins while making it an ideal candidate for others.

While phosphorylation is an area ripe for chemical discoveries, an even less understood form of protein regulation is the *cis-trans* isomerization of proline. Proline is unique among the common amino acids because its side chain is tied into the polypeptide backbone creating extreme restrictions in its conformational space[13]. While most amino acids are dominated by the *trans* state of its peptide bond, proline can accommodate appreciable percentages of both the *cis* and *trans* states[14]. This peculiarity has allowed some proteins to evolve in a way such that they will only recognize the *cis* or *trans* state adding a level of regulation built into the protein's substrate itself, a rare feat in the biochemical literature[15]. An example of this are the proline directed phophotases and kinases: proteins which phosphorylate and de-phosphorylate proteins, but only when the prolyl peptide bond is in the *trans* state [16, 17]. The *cis-trans* isomerization of

proline also plays a role in ion gated channels and is a rate limiting step in the protein folding process [14]. While proline provides nature with a convenient way to use conformational constraints to regulate proteins, it is not without its drawbacks. The primary disadvantage with this method of protein regulation is that prolyl *cis-trans* isomerization is slow on the biological time scale[14]. Typically, biological reactions need to occur on the millisecond timescale to be useful, but the *cis-trans* isomerization of proline typically occurs on the second to minute time scale[18]. Nature, however, has evolved the Peptidyl Prolyl Isomerases (PPIases) to enzymatically accelerate the *cis-trans* isomerization of proline to the millisecond timescale rendering it a useful way to regulate protein function[19].

The PPIases are proteins that are related by function and not structure[19]. Some examples include the cyclophilins, FKBP's, and the parvulins of which Pin1 is a member[19]. Cyclophilin A is a heavily studied PPIase because it has been shown that HIV protease actually uses the cyclophilin A present in the host organism to catalyze the prolyl *cis-trans* isomerization of a Gly-Pro sequence in order to replicate[20]. The FKBP's are known to be important in the prevention of the development of various neurological diseases[21]. It has been shown in knockout and mutation studies that decreased activity of the FKBP's allows diseases such as Alzheimer's to develop[22]. Pin1 is another PPIase that is important in neurological diseases and is known to regulate various transitions in mitosis[23]. It is also responsible for regulating the proline directed phosphotases and kinases[24]. These varied regulatory functions in the cell make the PPIases an ideal target for drug design, and a mechanistic understanding of their catalysis would greatly advance the field of medicine. The focus of the present work is Pin1 and the prolyl cis-trans isomerization of its protein substrates. In contrast to some of other PPIases, Pin1 is extremely strict in its substrate selection [25]. The catalytic domain of Pin1 will only bind proline residues when there is a

phosphorylated serine or phosphorylated threonine residue immediately preceding proline[25, 12]. While many of the diseases in which Pin1 plays a role are known in addition to many of its protein substrates, little is known about the atomistic basis for its catalysis. The purpose of this thesis is to elucidate details about Pin1's method of substrate selection and details about its mechanism of catalysis through the use of computational methods. Additionally, we explore some of the physical underpinnings of prolyl cis-trans isomerization in the case where a PPIase is absent for catalysis.

1.1 References

1. Crick, F.; Watson, J. D., Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953,** *171* (4356), 737-738.

2. Hao, Z.; Hong, S.; Chen, X.; Chen, P. R., Introducing Bioorthogonal Functionalities into Proteins in Living Cells. *Accounts of Chemical Research* **2011,** *44* (9), 742-751.

3. Seo, J.; Lee, K. J., Post-translational modifications and their biological functions: Proteomic analysis and systematic approaches. *Journal of Biochemistry and Molecular Biology* **2004,** *37* (1), 35-44.

4. Clayton, A. L.; Hazzalin, C. A.; Mahadevan, L. C., Enhanced Histone Acetylation and Transcription: A Dynamic Perspective. *Molecular Cell* **2006,** *23* (3), 289-296.

5. Choudhary, C.; Kumar, C.; Gnad, F.; Nielsen, M. L.; Rehman, M.; Walther, T. C.; Olsen, J. V.; Mann, M., Lysine Acetylation Targets Protein Complexes and Co-Regulates Major Cellular Functions. *Science* **2009,** *325* (5942), 834-840.

6. Cox, M.; Nelson, D. R.; Lehninger, A. L., *Lehninger Principles of Biochemistry*. W.H. Freeman: San Francisco, 2005.

7. Kukuruzinska, M. A.; Lennon, K., Protein N-Glycosylation: Molecular Genetics and Functional Significance. *Critical Reviews in Oral Biology & Medicine* **1998,** *9* (4), 415-448.

8. Vigerust, D., Protein glycosylation in infectious disease pathobiology and treatment. *Central European Journal of Biology* **2011,** *6* (5), 802-816.

9. Martin, L.; Latypova, X.; Terro, F., Post-translational modifications of tau protein: Implications for Alzheimer's disease. *Neurochemistry International* **2011,** *58* (4), 458-471.

10. Hamelberg, D.; Shen, T.; McCammon, J. A., Phosphorylation Effects on cis/trans Isomerization and the Backbone Conformation of Serine−Proline Motifs: Accelerated Molecular Dynamics Analysis. *Journal of the American Chemical Society* **2005,** *127* (6), 1969-1974.

11. Lee, Tae H.; Chen, C.-H.; Suizu, F.; Huang, P.; Schiene-Fischer, C.; Daum, S.; Zhang, Yan J.; Goate, A.; Chen, R.-H.; Zhou, Xiao Z.; Lu, Kun P., Death-Associated Protein Kinase 1 Phosphorylates Pin1 and Inhibits Its Prolyl Isomerase Activity and Cellular Function. *Molecular Cell* **2011,** *42* (2), 147-159.

12. Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J.-U.; Lu, K. P.; Fischer, G., Role of Phosphorylation in Determining the Backbone Dynamics of the Serine/Threonine-Proline Motif and Pin1 Substrate Recognition†. *Biochemistry* **1998,** *37* (16), 5566-5575.

13. Beck, D. A. C.; Alonso, D. O. V.; Inoyama, D.; Daggett, V., The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences* **2008,** *105* (34), 12259-12264.

14. Kun Ping, L.; Finn, G.; Tae Ho, L.; Nicholson, L. K., Prolyl cis-trans isomerization as a molecular timer. *Nature Chemical Biology* **2007,** *3* (10), 619-629.

15. Hamelberg, D.; McCammon, J. A., Mechanistic Insight into the Role of Transition-State Stabilization in Cyclophilin A. *Journal of the American Chemical Society* **2008,** *131* (1), 147-152.

16. Brown, N. R.; Noble, M. E.; Endicott, J. A.; Johnson, L. N., The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nature Cell Biology* **1999,** *1* (7), 438-443.

17. Weiwad, M.; Küllertz, G.; Schutkowski, M.; Fischer, G., Evidence that the substrate backbone conformation is critical to phosphorylation by p42 MAP kinase. *FEBS Letters* **2000,** *478* (1–2), 39-42.

18. Wolfenden, R.; Snider, M. J., The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Accounts of Chemical Research* **2001,** *34* (12), 938-945.

19. Göthel, S. F.; Marahiel, M. A., Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cellular and Molecular Life Sciences CMLS* **1999,** *55* (3), 423-436.

20. Yoo, S.; Myszka, D. G.; Yeh, C.-y.; McMurray, M.; Hill, C. P.; Sundquist, W. I., Molecular recognition in the HIV-1 capsid/cyclophilin A complex. *Journal of Molecular Biology* **1997,** *269* (5), 780-795.

21. Babine, R. E.; Villafranca, J. E.; Gold, B. G., FKBP immunophilin patents for neurological disorders. *Expert Opinion on Therapeutic Patents* **2005,** *15* (5), 555-573.

22. Cao, W.; Konsolaki, M., FKBP immunophilins and Alzheimer's disease: A chaperoned affair. *Journal of Biosciences* **2011,** *36* (3), 493-498.

23. Lu, K.; Hanes, S.; Hunter, T., A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **1996,** *380* (6574), 544-7.

24. Lu, K. P.; Zhou, X. Z., The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nature Reviews. Molecular Cell Biology* **2007,** *8* (11), 904-916.

25. Yaffe, M. B.; Schutkowski, M.; Shen, M.; Zhou, X. Z.; Stukenberg, P. T.; Rahfeld, J. U.; Xu, J.; Kuang, J.; Kirschner, M. W.; Fischer, G.; Cantley, L. C.; Lu, K. P., Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. *Science (New York, N.Y.)* **1997,** *278* (5345), 1957-1960.

2 CONFORMATIONAL SELECTION IN THE RECOGNITION OF PHOSPHORYLATED SUBSTRATES BY THE CATALYTIC DOMAIN OF PIN1

2.1 Abstract

Post-translational phosphorylation and the related conformational changes in signaling proteins are responsible for regulating a wide range of sub-cellular processes. Human Pin1 is central to many of these cell-signaling pathways in normal and aberrant sub-cellular processes, catalyzing *cis-trans* isomerization of the peptide ω-bond of phosphorylated serine/threonine-proline motifs in several proteins. Pin1 has therefore been identified as a possible drug target in many diseases, including Cancer and Alzheimer's. The effects of phosphorylation on Pin1 substrates, and the atomistic basis for Pin1 recognition and catalysis, are not well understood. Here, we determine the conformational consequences of phosphorylation on Pin1 substrate analogues and the mechanism of recognition by the catalytic domain of Pin1 using all-atom molecular dynamics simulations. We show that phosphorylation induces backbone conformational changes on the peptide substrate analogues. We also show that Pin1 recognizes specific conformations of its substrate by conformational selection. Furthermore, dynamical correlated motions in the free Pin1 enzyme are present in the enzyme of the enzyme-substrate complex when the substrate is in the transition state configuration, suggesting that these motions play significant roles during catalytic turnover. These results provide detailed atomistic picture of the mechanism of Pin1 recognition that can be exploited for drug design purposes and further our understanding of the synergistic complexities of posttranslational phosphorylation and *cis-trans* isomerization.

2.2 Introduction

Conformational transitions and the switching of biomolecules play important roles in sub-cellular processes. These processes generally involve recognition and transient interactions between biomolecules. The local switching of the isomeric state of the peptide bond preceding proline (also known as the ω-bond angle) from *trans* to *cis*, or vice versa, that brings about conformational changes in proteins are known to regulate a wide range of sub-cellular processes[1]. When the prolyl ω-bond is in the *trans* configuration, the protein may have a particular function that is altered when the prolyl ω-bond is switched to the *cis* configuration. Biological signaling processes utilize the additional conformational diversity as a result of the two possible isomeric states of the ω-bond angle. However, this signaling mechanism is very slow and is therefore regulated by ubiquitous peptidyl proline *cis-trans* isomerases (PPIases), speeding up *cis-trans* isomerization of the prolyl peptide bond by several orders of magnitude[2]. The role of *cis-trans* isomerization in Interleukin tyrosine kinase (Itk) SH2 domain[3] , ligand gated 5-HT3 ion channel[4,5], HIV capsid formation[6], and protein folding[7] are few familiar examples. Unlike the two well-known PPIases, cyclophilins[8] and FKBPs[9], the more recently discovered human Pin1 is a phosphorylation dependent PPIase[10]. Pin1 is specific for prolyl ω-bonds when the preceding residue is a phosphorylated serine (pSer) or phosphorylated threonine (pThr), combining the regulatory role of post-translational phosphorylation and conformational switching of the prolyl ω-bonds (Figure 1). In particular, Pin1 is involved in mitosis, various cancers, and Alzheimer's disease[11,12]. Additionally, many important proteins in humans have been identified as substrates for Pin1[13], including RNA polymerase II[14].

The discovery of Pin1 has provided additional evidence for the involvement of PPIases in cell signaling[15]. Pin1 plays an important role in the G_2-M phase transition in mitosis through its interaction with Topoisomerase[11]. This interaction allows Pin1 to localize chromatin in the G_2-M phase of mitosis. Pin1 also interacts with many other important mitotic proteins such as CDC25C, EM11, and WEE1[15]. Inhibition or deletion of Pin1 leads to mitotic catastrophe and chromatin condensation, therefore Pin1 is a key enzyme in mitosis. During the G_1-S phase transition, Pin1 interacts with Cyclin D1, creating stabilization and nuclear localization for the transition, and the level of Pin1 in normal cells is elevated because many of its protein substrates control this transition. In various cancers (breast, thyroid, and prostate), Pin1 and its substrate Cyclin D1 are over-expressed, indicating that both play an important role in oncogenesis[13]. Consequently, Pin1 has been an effective early diagnostic marker in prostate cancer[16]. Pin1 is viewed as a "double-edged sword."[12] In normal cells, phosphorylation of Ser/Thr-Pro motifs is involved in many signaling pathways. Similarly, in cancer cells, phosphorylation of Ser/Thr-Pro motifs follows oncogenic signaling. As a result, Pin1 regulates proteins that are involved in both promotion and suppression of oncogenesis and has been identified as a target for drug design[16]. In Alzheimer's disease, it is believed that Pin1 actually protects against the disease, since it is deactivated in the early stages of the illness through oxidative modification[11]. More specifically, phosphorylation and dephosphorylation of the tau protein regulate the assembly of microtubules in neuronal cells, and Pin1 regulates *cis-trans* isomerization of numerous pSer/pThr-Pro motifs in tau protien[17]. When Pin1 is depleted, tangling of the hyperphosphorylated tau takes place, which is one of the hallmarks of Alzheimer's disease, eventually resulting in cell death[18]. Clearly, deciphering the exact role of Pin1 and its mechanism of action at the atomistic level will

provide a deeper understanding of many diseases and will help in the design of new classes of drug candidates.

The distinguishing characteristic of Pin1 as compared to the other PPIases is that it is selective for peptide ω-bond of pSer/pThr-Pro motifs. Therefore, phosphorylation of Ser/Thr-Pro motifs facilitates the interaction between Pin1 and its protein substrates. It is known that phosphorylation perturbs the free energy landscape of proteins[19-21]. However, the effects of phosphorylation on the conformational landscape of Pin1 substrates and how phosphorylation facilitates recognition by Pin1 is not well understood. In addition, x-ray crystal structures of Pin1 bound to its natural substrates have not been solved to date. Therefore, a detailed atomistic understanding of how Pin1 interacts with its substrate is lacking and will aid in the development of new classes of inhibitors. Also, knowledge of the interactions of Pin1 with its natural substrates could advance our understanding of the mechanism of *cis-trans* isomerization of Pin1 and possibly of other isomerases, a mechanism that is still not well understood and controversial[2,22]. In understanding the effects of phosphorylation on the conformation of Pin1 substrates and the role phosphorylation plays in the recognition by Pin1, we focus here on studying the effects of phosphorylation on the conformational preferences of well-studied substrate analogues, Ace-Ala-Ala-X-Pro-Phe-Nme, where X is pSer or pThr, and their interactions with Pin1 using all-atom molecular dynamics (MD) simulations in explicit water. Simulations of the free substrate analogues and one of the enzyme-substrate complexes were carried out using accelerated molecular dynamics (aMD)[23], which allows for a very efficient and thorough sampling of the configuration space and the slow *cis-trans* isomerization of the prolyl ω-bond.

2.2 Methods

All simulations were carried out using the Amber 10 suite of programs[54] and the modified parm99SB[55] version of the Cornell el al.[56] force field, with the re-optimized dihedral parameters for the peptide backbone ω-bond[57]. The NPT ensemble at a pressure of 1 bar and a temperature of 300K was used for all the simulations with explicit TIP3P[58] water model, and the SHAKE[59] algorithm was used to constrain all bonds involving hydrogen. The Langevin thermostat was used for regulating the temperature of the system to 300 K with a collision frequency of 1.0 ps^{-1}. Short-range nonbonded interactions were calculated with a cutoff of 9 Å, and all long range interactions were calculated using the particle mesh Ewald summation[60]. A timestep of 2 fs was used to integrate Newton's equations of motions.

2.2.1 Accelerated Molecular Dynamics of Pin1 substrate analogue

The substrate analogues, Ace-Ala-Ala-X-Pro-Phe-Nme, were build with the XLEAP program, where X was Ser, Thr, pSer, and pThr. The parameters used for the pSer and pThr residues were developed by Homeyer et al.[61] The dianionic form of pSer and pThr was used in this study because it was shown that at around neutral pH the phosphate group is deprotonated and Pin1 prefers to bind its substrates in the dianionic form even at low pH[33,43]. The substrates were placed in a cubic periodic water box with the edges of the box at least 10 Å away from any part of the substrates. Two Na+ ions were added to the systems of phosphorylated substrates in order to attained electrostatic neutrality. The systems were then equilibrated with a series of minimization and molecular dynamics simulations. Each equilibrated systems was then simulated for 260 ns using accelerated molecular dynamics. A boost energy, E, of 120 kcal/mol and α of 20 kcal/mol were used for the acceleration. The accelerated molecular dynamics was implemented in a modified version of pmemd in the Amber 10 suite of programs. Each

simulation was then repeated 10 times. A snapshot of the simulation was written to the trajectory file every 5 steps. The high frequency of data collection was used in order to reduce errors due to reweighting[62]. Ptraj and Matlab were used to analyze the data.

2.2.2 Molecular Dynamics Simulations of Pin1 and Pin1 substrate complexes

The starting structure for the current study was a 1.7 Å resolution x-ray crystal structure of an inhibitor bound complex of Pin1 with PDB ID 2Q5A[35]. Only the catalytic domain was used. The WW domain was remove, since it does not affect the catalytic activating and binding of the catalytic domain[43]. The inhibitor was removed to prepare the system for the free Pin1 enzyme. In order to prepare the enzyme-substrate bound complexes, the peptidomimic inhibitor was modified to the sequence of the substrate analogues. The peptidomimic inhibitor already provided the peptide backbone template, and the side chains of the non-natural amino acids were modified to amino acids in our sequence by retaining the common atoms. Xleap was used to add the missing atoms. The system were solvated in a periodic octahedron water box with the edges of the box at least 10 Å away from any part of the solute. The systems were then neutral with counter ions. The systems were equilibrated with a series of minimization and molecular dynamics simulations, initially applying a harmonic constraint only on the solute with a force constant of 200 kcal/mol/Å^2 and reducing it by half for each subsequent equilibration step until it is 25 kcal/mol/Å^2. A final molecular dynamics simulation was carried out without any harmonic constraints. The systems were then simulated for 110ns each. The first 10ns was discarded as additional equilibration. The resulting conformation of the peptide ω-bond of the X-Pro motif of the substrates was in the *cis* configuration. The principal component analysis was carried out by combining the trajectories of free Pin1 and Pin1 in the different states of the substrates (*cis, trans,* and transition state) using the Ptraj program.

2.3 Results and Discussion

We have carried out extensive sampling of the conformational space of Pin1 substrate analogues, Pin1-substrate complexes, and free Pin1, using all-atom molecular dynamics simulations in explicit water. These simulations have allowed us to study the effects of phosphorylation on the conformations of the substrates, containing the pSer/pThr-Pro motifs, and the mechanism of recognition by the catalytic domain of Pin1. Posttranslational modification of proteins by phosphorylation is intricately involved in many normal and aberrant signaling pathways in biology. Therefore, understanding the exact role of phosphorylation in biology is a key to treating many diseases.

2.3.1 Phosphorylation effects on the equilibriumof Pin1 substrate analogues

Accelerated molecular dynamics simulations were carried out on Pin1 substrate analogues (Ace-Ala-Ala-X-Pro-Phe-Nme), where X was Ser, Thr, pSer, or pThr. Each substrate was simulated for 260 ns, and each simulation was repeated 10 times for a total of 2.6 μs. The results of conformational sampling of the Ser/pSer-Pro motifs of the free substrates are summarized in Figures 2 and 3. The results for Thr/pThr-Pro are similar to those of Ser/pSer-Pro and are summarized in Figures S1 and S2. The Ramachandran or φ/ψ plots for X = Ser and Thr reveal that the right handed α-helical region (φ ~ -75˚, ψ ~ -50˚) for serine is slightly more populated than that for threonine. The left handed helical regions for both X = Ser and Thr are reasonably populated, although the left-handed basin for X = Ser is slightly broader. Nonetheless, the β-region for both X = Ser and X = Thr is the most populated, with a very broad basin. These results are similar to those obtained by Lee et al.[24] in their study of the Gly-Ser-Ser-Ser peptide, but with some differences. Their study concluded that prior to phosphorylation the primary configuration of serine is a mixture of β and polyproline II conformations, with very

little or no α-helical conformation. In the current study, significant α-helical conformation is seen prior to phosphorylation, but the β-conformation still dominates.

A previous NMR study by Tholey et al.[25] showed that phosphorylation of the Gly-Ser-Ser-Ser peptide results in changes to the ^{3}J coupling constants of the backbone angles, indicating a change in the equilibrium conformation of the peptide. These results were later supported by Shen et al.[26] using atomistic Brownian dynamics simulations, showing that phosphorylation increases the α-helical population of the peptide backbone conformation. Similar results were also obtained for the Thr-pSer-Pro-Ile peptide as well as for the pSer/pThr-Pro dipeptide[27,28]. A similar effect is also observed here for Pin1 substrate analogues. Upon phosphorylation, i.e. when X = pSer or pThr, there are two main observations: a considerable increase in the population of the right-handed α-helical conformation and a decrease in the population of the left-handed helix for both the pSer and pThr containing substrates as compared to the unphosphorylated substrates. The population of the β-region in both substrates, however, does not change significantly upon phosphorylation (Figures 2B and S1B). In the present study, the majority of the increase in the population of the α-helical region has been at the expense of the left handed region. The population of the left-handed helical region of the X = pThr is almost wiped out upon phosphorylation, while a small population of left handed helical formation is maintained when X = pSer. Also, phosphorylation decreases the left-handed helical region of the flanking residues of X-Pro, but does not significantly alter their α-β equilibrium, as summarized in Figures S3 and S4. Therefore, the results suggest that the effects of phosphorylation on the backbone conformation of the substrate can propagate beyond the phosphorylation site. One major difference between the two types of substrates is that the high energy or “forbidden” region (energies of more than 40 kcal/mol) is larger in the cases when X = Thr and pThr than

when X = Ser and pSer, respectively. Therefore, the Thr/pThr-Pro motif seems to be less flexible than the Ser/pSer-Pro motif. Beck et al.[29] have shown that threonine containing motifs covers less space on the Ramachandran plot than serine containing motifs. Our results show that this is also true for the phosphorylated motifs, and these subtle differences between the conformations of pThr-Pro and pSer-Pro motifs and their effects on the substrate analogues of Pin1 could have implications for the slight differences in their recognition and catalysis by Pin1. For example, it appears that pThr-Pro containing substrates binds better to Pin1 than that of pSer-Pro[30], but the catalytic activity of Pin1 for pSer-Pro containing substrate seems higher than that of pThr-Pro substrates[31].

Previously, it was shown that there is a dependency of ω-bond angle on the backbone ψ angle of Pro for X-Pro motifs[27,32]. It turns out that there is also a dependency of the ω-bond angle on the backbone ψ angle of X. In Figures 2 and 3 (bottom), we have plotted the ψ-angle of Ser and pSer as a function of the ω-bond angle and that of Pro for the unphosphorylated and phosphorylated substrate analogues, respectively. The equivalent plots for Thr/pThr-Pro motifs are shown in Figures S1 and S2. As expected, the predominant isomer for the ω-bond angle is *trans* for all the substrates, as can be seen from the ω-ψ spaces in Figures 2 and 3. Interestingly, the α-helical and β regions defined by the ψ angle of X and Pro can easily and are almost equally populated when the ω-bond angle is in the *trans* configuration for the unphosphorylated substrates. However, the β region of the ψ angle of X and Pro is more populated when the ω-bond angle is in the *cis* configuration for the unphosphorylated substrates as compared to the α-helical region. Also, the barrier separating the α-helical and β regions is smaller when the ω-bond angle is in the *trans* configuration, as compared to when it is in the *cis* configuration, therefore it is easier for X-Pro motifs to change backbone conformations when they are in the *trans*

configuration. The barrier separating the α-helical and β regions, when the ω-bond angle is in the *trans* configuration, is slightly lower for Pro than for X.

Upon phosphorylation, there is a slight reduction in the barrier separating the α-helical and β regions for Pro when the ω-bond angle is in the *cis* configuration, creating a thermodynamically stable valley for the conformational transition. The same barrier for X did not change much upon phosphorylation. Therefore, the results suggest that phosphorylation can act to also facilitate the conformational transition between the α-helical and β regions for Pro when the ω-bond angle is in the *cis* state. In addition to lowering the barrier between the α-helical and β regions of Pro when the ω-bond is in the *cis* configuration, phosphorylation also increases the α-helical population of Pro. The change in population is more pronounced when the ω-bond is in the *cis* configuration, as shown in Figure 3.

2.3.2 Conformational Selection in Pin1 Recognition

The overall *cis-trans* equilibrium and barrier along the ω-bond are only slightly affected by phosphorylation. The barrier from the *trans* configuration to the *cis* configuration is calculated to be ~20 kcal/mol for both the unphosphorylated and phosphorylated substrate analogues, very similar to experimentally measured values[33]. Since phosphorylation does not seem to greatly affect the *cis-trans* barrier, it can be argued that phosphorylation renders the Ser/Thr-Pro motifs almost unrecognizable by the other ubiquitous and promiscuous isomerases (cyclophilins and FKBPs), and specific for Pin1 recognition. What role does phosphorylation therefore play in Pin1 recognition and specificity?

X-ray crystal structures of natural substrates or substrate analogues bound to Pin1 have not been solved to date. However, structure based drug design has provided several lead compounds to target Pin1[34], and x-ray crystal structures for some of those compounds in the

active site of Pin1 have been solved[35-37]. In order to study the mechanism of Pin1 recognition of its natural substrates, we have modeled the enzyme-substrate complex from a 1.7 Å resolution x-ray crystal structure of a phosphorylated peptidomimetic inhibitor bound in the active site of Pin1[35], with PDB ID 2Q5A, as shown in Figure 4A. The active site residues of Pin1 are shown in Figure 4B. The peptidomimetic inhibitor serves as an ideal starting point for modeling the natural substrate because it readily provides the template for the peptide backbone, and non-natural amino acid side chains can be easily modified to natural amino acids by keeping the common atoms. We modified the Pin1-bound inhibitor to the substrate analogues with X = pSer or X= pThr. The complexes were then solvated with explicit water, minimized and equilibrated for 10 ns. The final equilibrated structure of the complex between Pin1 and the substrate analogue when X = pSer is shown in Figure 4D. The active site can be broadly separated into two compartments. His 59, Leu 61, Cys 113, Leu 122, Met 130, Gln 131, Phe 134, Thr 152, Ser 154, and His 157 (residues on the left of Figure 4B), form the X-Pro binding pocket that mainly recognizes the side chain of proline. Lys 68, Arg 68, and Arg 69 form the binding pocket for the phosphate moiety of pSer that could form a claw-like structure around the phosphate moiety[38]. The configuration of the resulting X-Pro ω-bond angle is in the *cis* isomer ($\omega \sim 0^o$) after the modification. Simulations of both complexes with the ω-bond angle of the X-Pro motif of the substrate analogues in the *cis* conformation were each extended for 100 ns. The ω-bond angle of the X-Pro motif of the substrates for both complexes was gradually converted to the *trans* (180^o) conformation over 1 ns MD simulation using angle restraints, and the complexes were allowed to relax with the ω-bond angle of the substrate in the *trans* conformation for an additional 10 ns. The *trans* complexes were also then simulated for an additional 100 ns.

Analyses of the backbone dihedral angles of the substrate when bound to Pin1 suggest that, in addition to recognizing the phosphate group of the substrate, Pin1 prefers specific backbone conformations of the substrate depending on the configuration of the ω-bond angle. When the ω-bond angle of the substrate is in the *cis* configuration, the backbone conformations of the X-Pro motifs are in the β region, and this is true for both X = pSer and X = pThr as shown in Figures 5 and S5. However, when the ω-bond angle of the substrates is in the *trans* configuration, the backbone conformations of the X-Pro motif of the substrates in the complexes are mainly in the α-helical region (Figures 5 and S5). The backbone conformation of the substrates is localized in the complex and is similarly for both pSer and pThr containing substrates. The results therefore suggest that Pin1 recognizes its two different substrates in a similar fashion and selects out a very small sub-states of the many different conformations that are accessible by the free substrates. It is interesting to note that the backbone conformation of Pro, when the ω-bond angle is in the *cis* configuration for both free substrates (when X = pSer and pThr) is also mainly in the β region, and Pin1 recognizes this dominant conformation in the *cis* configuration. We also carried out accelerated molecular dynamics of the Pin1-substrate complex, when X = pSer, for 260 ns, in order to study cis-trans isomerization of the X-Pro ω-bond in the active site of Pin1. Similar to the results obtained using normal molecular dynamics simulation, Pin1 prefers the substrate in the α-helical conformation when the ω-bond angle is in the *trans* configuration and the β region when the ω-bond angle is in the *cis* configuration, as shown in Figure 6. The results further suggest that the ease of *cis-trans* isomerization along the ω-bond angle depends on the backbone conformation of Pro. The *cis-trans* barrier height on the positive side of the ω-bond angle, when the backbone conformation of Pro is α-helical, is lower than the other possible paths, as shown in Figure 6. These results therefore present compelling

evidence that Pin1 recognition and binding, and possibly catalysis, rely on more than just the electrostatic effects of the phosphate moiety, but also on the ensuing changes in the secondary structure of the substrates. This interplay between the backbone conformations and isomeric states of the Pin1 substrates and Pin1 recognition could be very important for proper function. For example, proline-directed serine/threonine kinases and phosphotases phosphorylate and dephosphorylate Ser/Thr-Pro motifs, respectively, only if the ω-bond angle is in the *trans* configuration[17,39,40]. Therefore, Pin1 is critical to maintaining the *cis-trans* equilibrium and regulating the function of its substrates and other proteins.

Interestingly, in addition to recognizing specific conformations of the substrate analogues in the different states, the mechanism of Pin1 recognition is also based on conformational selection, which is believed to be a widespread mechanism in protein recognition[41,42]. We carried out principal component (PCA) analysis on the simulations of free Pin1 and the substrate-bound complexes of Pin1, when the ω-bond angle of the substrate is in *trans*, *cis* and transition state configurations (when X = pSer). The simulations of the transition state of the substrate analogues bound to Pin1 were also carried out for 100 ns, using normal molecular dynamics, while keeping the ω-bond angle at around 90^{o} with an angle force constant restraint of 1000 kcal/mol/rad^2. Similar results were obtained for X = pThr, therefore only the results for X = pSer are shown. Figure 7 shows the distribution of the first two eigenvectors of the slowest principal components (PC) projected back on the trajectories. The eigenvalues of the top five eigenvectors are 197.2, 23.2, 12.8, 11.5 and 8.9 Å^2, therefore, it can be seen that the first two PCs characterize almost 90% of the motions described by the top five PCs, and the first PC represents more than 70% of all the motions. Figure 7A shows the distribution of PC1 of Pin1 in the different states. Free Pin1 spans a large region of PC1 with approximately three peaks, representing the different

conformations of Pin1 along PC1. However, it is interesting to see that the conformations of Pin1 that bind the substrate analogues when they are in the *cis*, *trans*, and *transition state* configurations are subsets of the free enzyme. The conformations of Pin1 that bind the *cis* and *trans* configurations of the substrate are similar, representing the least populated peak of free Pin1. The ensemble of conformations that binds the *trans* configuration has a narrower distributions than that of the *cis* configuration. However, the *transition state* configuration is recognized by a broader range of conformations that partly include that of the *cis* and *trans* configurations and, to a larger extent, one of the two dominant ensembles of conformations of the free enzyme. Therefore distribution of the conformations along PC1 that bind the transition state is much broader than that of the *cis* and *trans* configurations, as shown in Figure 7.

The motion of Pin1 that is most dominant and captured by the first principal component is also shown as insets in Figure 7. This motion represents a hinge-like motion between an α-helix and a β-turn, with the proline binding pocket sitting in-between the two. Therefore, this hinge-like motion modulates the active site and could possibly have a meaningful role to play in the catalytic mechanism of the enzyme. The different ensembles of conformations that recognize the different configurations of the substrate analogue make slightly different interactions with the pSer-Pro motif. The side chain of Pro of the substrate forms a hydrophobic interaction with the proline binding pocket of Pin1 (Figure 4) when the substrate is in the *cis*, *trans* and transition state configurations. Also the phosphate moiety of the side chain of pSer always forms an electrostatic interaction with the phosphate binding pocket (also Figure 4) irrespective of the state of the substrate. Arg69 and Lys63 form more stable interactions with the phosphate moiety than Arg68, as shown in Figure S6. Arg69 and Lys63 are more localized around the phosphate group, while Arg68 is more mobile and rarely interacts with the phosphate group. Unlike the *cis*

and *trans* states, the results suggest that the transition state of the substrate interacts a little more favorably with Arg68. These results agree well with earlier mutational studies showing that Arg69 is more important for binding than Arg68[43].

2.3.3 Dynamical and Correlated Motions in Pin1

The motion of Pin1 along the first principal component is shown to be the most dominant (Figure 7) and spans a large range in the free enzyme. This motion is dominant in the enzyme of the *transition state* bound complex and also spans a relatively large range, as compared to that of the *trans* and *cis* bound complexes. This led us to suggest that in addition to the ability of the enzyme to stabilize the transition state during catalysis this dominant motion in the free enzyme could play key roles during catalytic turnover. This is interesting because the well-studies cyclophilin A, which is a *cis-trans* isomerase of a different family, has been shown to possess motions in the free enzyme that are present in the enzyme-substrate complex during catalytic turnover. These motions are believed to assist in catalysis by moving the substrate from one state to the other[44-46]. This dynamical phenomenon has also been observed in other enzymes[47,48], including Pin1[49-51]. However, the effect of these dynamical motions to the catalytic rate and whether they provide any enhancement to the rate of catalysis as compared to the isomerization of the substrate (free in solution) is not clear and still being debated[52,53]. What is clear is that these dynamical motions are generally coupled with the catalytic process, as the enzymes are very dynamic and have to alter their conformations to recognize the substrate in the reactant, transition, and product states, as observed in this work.

In addition to the dynamical motions represented by the first principal component, we have also calculated the dynamical correlated motions in free Pin1 and the substrate bound complexes in the three different states, as shown in the dynamical cross correlation map in

Figure 9. It is interesting and surprising to see that correlated motions in the free enzyme (labeled as *m1*, *m2*, and *m3* in Figure 9) are also present in the enzyme conformations that bind the *transition state*. However, they are not present in the *trans*-enzyme complex and are only slightly present in the *cis*-enzyme complex. The correlated motions in free Pin1 are very similar to that of the *transition state* complex than the *cis* and *trans* complexes (Figure 9). All of the highlighted motions involve amino acids between residues 122 to 140 (shown as orange in Figure 9) and three other parts of the enzyme (shown as yellow in Figure 9). All of the correlated motions in one way or the other involve residues that make up the active site of Pin1. Residues 122 to 140 contain most of the residues that make up the proline binding pocket, including Phe134 that is at the base of the cavity. The *m1* correlated motion is between the majority of the proline binding site residues (Leu122, Met130, Glu131, and Phe134) and three other active site residues, Thr152, Ser154, and His157, one of which (Ser154) is actively involved in hydrogen bonding with the substrate when the substrate is in the *trans* and transition state configurations. The correlated motion depicted by *m2* is again between the majority of the proline binding site residues (Leu122, Met130, Glu131, and Phe134) and two other active site residues, Cys113 and Ser114. Cys113 can hydrogen bond with the substrate when it is in the *cis* configuration and also when it is in the *transition state*. The third correlated motion, *m3*, involves the proline binding pocket and the residues that make up the phosphate binding pocket (Lys63, Arg68 and Arg69). Several of these residues have previously been identified as being important for catalysis using NMR relaxation experiments[49-51]. Our results therefore suggest that these motions are relevant to catalysis, since they involve all of the residues that make up the active site of the enzyme. Also, the correlated motions in the free enzyme show up only slightly when the substrate is in the *cis* configuration. The correlated motions become intense when the substrate is in *the transition state*

configuration and are not present at all when the substrate is in the *trans* configuration. Therefore, these motions are coupled with the catalytic process in moving the substrate across the transition state from the *cis* to *trans* configurations and vice versa.

2.4 Concluding Remarks

In this study, we have shown that phosphorylation is an important mechanism by which transient conformational changes can be induced on proteins in post-translational modification of serine or threonine. The phosphate moiety and perturbation of the equilibrium distribution of protein conformations control the recognition mechanism with other proteins, such as human Pin1, and are central to regulating many intricate sub-cellular processes. Our results suggest that Pin1 recognizes specific backbone conformation of its substrates using conformational selection and has a wider range of conformations that recognize the transition state. These different ensembles of conformations of Pin1 can be separately targeted with slightly different classes of compounds as a drug designing strategy, exploiting the conformational variability of Pin1. Also, correlated motions in the free Pin1 are present in the ensemble of conformations that bind the substrate in the *transition state* configuration. The exact role of these dynamical coupled motions in the catalytic activity of Pin1 is yet to be determined and requires much further study. Fully understanding the atomistic basis for Pin1 recognition and mechanism of the catalysis of the *cis-trans* isomerization of the pSer/pThr-Pro motifs in signaling proteins will provide the foundation for the development of new classes of compounds to treat many diseases, including cancer and Alzheimer's.

2.5 Supporting Information

Six figures are included as supporting information. This material is available free of charge via the Internet at http://pubs.acs.org/doi/suppl/10.1021/bi2009954

2.6 Acknowledgments

We would like to thank Dr. Urmi Doshi for helpful discussions during the preparation of this manuscript.

2.7 References

1. Kun Ping, L., Finn, G., Tae Ho, L., and Nicholson, L. K. (2007) Prolyl cis-trans isomerization as a molecular timer, *Nature Chemical Biology 3*, 619-629.
2. Fanghanel, J., and Fischer, G. (2004) Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases, *Front Biosci 9*, 3453-3478.
3. Severin, A., Joseph, R. E., Boyken, S., Fulton, D. B., and Andreotti, A. H. (2009) Proline Isomerization Preorganizes the Itk SH2 Domain for Binding to the Itk SH3 Domain, *Journal of Molecular Biology 387*, 726-743.
4. Lummis, S. C., Beene, D. L., Lee, L. W., Lester, H. A., Broadhurst, R. W., and Dougherty, D. A. (2005) Cis-trans isomerization at a proline opens the pore of a neurotransmitter-gated ion channel, *Nature 438*, 248-252.
5. Melis, C., Bussi, G., Lummis, S. C. R., and Molteni, C. (2009) Trans−cis Switching Mechanisms in Proline Analogues and Their Relevance for the Gating of the 5-HT3 Receptor, *The Journal of Physical Chemistry B 113*, 12148-12153.
6. Yoo, S., Myszka, D. G., Yeh, C.-y., McMurray, M., Hill, C. P., and Sundquist, W. I. (1997) Molecular recognition in the HIV-1 capsid/cyclophilin A complex, *Journal of Molecular Biology 269*, 780-795.
7. Brandts, J. F., Halvorson, H. R., and Brennan, M. (1975) Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues, *Biochemistry 14*, 4953-4963.

8. Handschumacher, R., Harding, M., Rice, J., Drugge, R., and Speicher, D. (1984) Cyclophilin: a specific cytosolic binding protein for cyclosporin A, *Science 226*, 544-547.

9. Siekierka, J. J., Hung, S. H. Y., Poe, M., Lin, C. S., and Sigal, N. H. (1989) A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin, *Nature 341*, 755-757.

10. Lu, K., Hanes, S., and Hunter, T. (1996) A human peptidyl-prolyl isomerase essential for regulation of mitosis, *Nature 380*, 544-547.

11. Lu, K. P., and Zhou, X. Z. (2007) The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease, *Nature Reviews. Molecular Cell Biology 8*, 904-916.

12. Theuerkorn, M., Fischer, G., and Schiene-Fischer, C. (2011) Prolyl cis/trans isomerase signalling pathways in cancer, *Current opinion in pharmacology*, in press.

13. Ryo, A., Liou, Y.-C., Lu, K. P., and Wulf, G. (2003) Prolyl isomerase Pin1: a catalyst for oncogenesis and a potential therapeutic target in cancer, *Journal of Cell Science 116*, 773-783.

14. Shaw, P. E. (2007) Peptidyl-prolyl cis/trans isomerases and transcription: is there a twist in the tail?, *EMBO Reports 8*, 40-45.

15. Zhou, X. Z., Lu, P. J., Wulf, G., and Lu, K. P. (1999) Phosphorylation-dependent prolyl isomerization: a novel signaling regulatory mechanism, *Cellular and Molecular Life Sciences 56*, 788-806.

16. Finn, G., and Lu, K. P. (2008) Phosphorylation-specific prolyl isomerase Pin1 as a new diagnostic and therapeutic target for cancer, *Current Cancer Drug Targets 8*, 223-229.

17. Zhou, X. Z., Kops, O., Werner, A., Lu, P. J., Shen, M., Stoller, G., Kullertz, G., Stark, M., Fischer, G., and Lu, K. P. (2000) Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins, *Molecular cell 6*, 873-883.

18. Pei-Jung, L., Wulf, G., Zhou, X., Davies, P., and Lu, K. (1999) The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein, *Nature 399*, 784-788.

19. Shen, T., Zong, C., Hamelberg, D., McCammon, J. A., and Wolynes, P. G. (2005) The folding energy landscape and phosphorylation: modeling the conformational switch of the NFAT regulatory domain, *Faseb J 19*, 1389-1395.

20. Groban, E. S., Narayanan, A., and Jacobson, M. P. (2006) Conformational changes in protein loops and helices induced by post-translational phosphorylation, *PLoS computational biology 2*, e32.

21. Narayanan, A., and Jacobson, M. P. (2009) Computational studies of protein regulation by post-translational phosphorylation, *Current opinion in structural biology 19*, 156-163.

22. Hamelberg, D., and McCammon, J. A. (2009) Mechanistic insight into the role of transition-state stabilization in cyclophilin A, *J Am Chem Soc 131*, 147-152.

23. Hamelberg, D., Mongan, J., and McCammon, J. A. (2004) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules, *Journal of Chemical Physics 120*, 11919-11929.

24. Kyung-Koo, L., Cheonik, J., Seongeun, Y., Hogyu, H., and Minhaeng, C. (2007) Phosphorylation effect on the GSSS peptide conformation in water: Infrared, vibrational circular dichroism, and circular dichroism experiments and comparisons with molecular dynamics simulations, *Journal of Chemical Physics 126*, 235102.

25. Tholey, A., Lindemann, A., Kinzel, V., and Reed, J. (1999) Direct Effects of Phosphorylation on the Preferred Backbone Conformation of Peptides: A Nuclear Magnetic Resonance Study, *Biophysical Journal 76*, 76-87.

26. Shen, T., Wong, C. F., and McCammon, J. A. (2001) Atomistic Brownian Dynamics Simulation of Peptide Phosphorylation, *Journal of the American Chemical Society 123*, 9107-9111.

27. Hamelberg, D., Shen, T., and McCammon, J. A. (2005) Phosphorylation Effects on cis/trans Isomerization and the Backbone Conformation of Serine−Proline Motifs: Accelerated Molecular Dynamics Analysis, *Journal of the American Chemical Society 127*, 1969-1974.

28. Byun, B. J., and Kang, Y. K. (2010) Conformational preferences and prolyl cis–trans isomerization of phosphorylated Ser/Thr-Pro motifs, *Biopolymers 93*, 330-339.

29. Beck, D. A. C., Alonso, D. O. V., Inoyama, D., and Daggett, V. (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins, *Proceedings of the National Academy of Sciences 105*, 12259-12264.

30. Zhou, W., Yang, Q., Low, C. B., Karthik, B. C., Wang, Y., Ryo, A., Yao, S. Q., Yang, D., and Liou, Y.-C. (2009) Pin1 Catalyzes Conformational Changes of Thr-187 in p27Kip1 and Mediates Its Stability through a Polyubiquitination Process, *Journal of Biological Chemistry 284*, 23980-23988.

31. Yaffe, M. B., Schutkowski, M., Shen, M., Zhou, X. Z., Stukenberg, P. T., Rahfeld, J. U., Xu, J., Kuang, J., Kirschner, M. W., Fischer, G., Cantley, L. C., and Lu, K. P. (1997)

Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism, *Science (New York, N.Y.) 278*, 1957-1960.

32. Fischer, S., Dunbrack, R. L., and Karplus, M. (1994) Cis-Trans Imide Isomerization of the Proline Dipeptide, *Journal of the American Chemical Society 116*, 11931-11937.
33. Schutkowski, M., Bernhardt, A., Zhou, X. Z., Shen, M., Reimer, U., Rahfeld, J.-U., Lu, K. P., and Fischer, G. (1998) Role of Phosphorylation in Determining the Backbone Dynamics of the Serine/Threonine-Proline Motif and Pin1 Substrate Recognition†, *Biochemistry 37*, 5566-5575.
34. Schiene-Fischer, C., Aumuller, T., and Fischer, G. Peptide Bond cis/trans Isomerases: A Biocatalysis Perspective of Conformational Dynamics in Proteins, *Topics in current chemistry*.
35. Zhang, Y., Daum, S., Wildemann, D., Zhou, X. Z., Verdecia, M. A., Bowman, M. E., Lücke, C., Hunter, T., Lu, K.-P., Fischer, G., and Noel, J. P. (2007) Structural Basis for High-Affinity Peptide Inhibition of Human Pin1, *ACS Chemical Biology 2*, 320-328.
36. Guo, C., Hou, X., Dong, L., Dagostino, E., Greasley, S., Ferre, R., Marakovits, J., Johnson, M. C., Matthews, D., Mroczkowski, B., Parge, H., Vanarsdale, T., Popoff, I., Piraino, J., Margosiak, S., Thomson, J., Los, G., and Murray, B. W. (2009) Structure-based design of novel human Pin1 inhibitors (I), *Bioorganic & medicinal chemistry letters 19*, 5613-5616.
37. Potter, A., Oldfield, V., Nunns, C., Fromont, C., Ray, S., Northfield, C. J., Bryant, C. J., Scrace, S. F., Robinson, D., Matossova, N., Baker, L., Dokurno, P., Surgenor, A. E., Davis, B., Richardson, C. M., Murray, J. B., and Moore, J. D. (2010) Discovery of cell-

active phenyl-imidazole Pin1 inhibitors by structure-guided fragment evolution, *Bioorganic & medicinal chemistry letters 20*, 6483-6488.

38. Hamelberg, D., Shen, T., and McCammon, J. A. (2007) A proposed signaling motif for nuclear import in mRNA processing via the formation of arginine claw, *Proc Natl Acad Sci U S A 104*, 14947-14951.

39. Brown, N. R., Noble, M. E., Endicott, J. A., and Johnson, L. N. (1999) The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases, *Nature cell biology 1*, 438-443.

40. Weiwad, M., Kullertz, G., Schutkowski, M., and Fischer, G. (2000) Evidence that the substrate backbone conformation is critical to phosphorylation by p42 MAP kinase, *FEBS letters 478*, 39-42.

41. Ma, B., and Nussinov, R. (2010) Enzyme dynamics point to stepwise conformational selection in catalysis, *Current opinion in chemical biology 14*, 652-659.

42. Grant, B. J., McCammon, J. A., and Gorfe, A. A. (2010) Conformational selection in G-proteins: lessons from Ras and Rho, *Biophys J 99*, L87-89.

43. Daum, S., Fanghanel, J., Wildemann, D., and Schiene-Fischer, C. (2006) Thermodynamics of phosphopeptide binding to the human peptidyl prolyl cis/trans isomerase Pin1, *Biochemistry 45*, 12125-12135.

44. Eisenmesser, E. Z., Millet, O., Labeikovsky, W., Korzhnev, D. M., Wolf-Watz, M., Bosco, D. A., Skalicky, J. J., Kay, L. E., and Kern, D. (2005) Intrinsic dynamics of an enzyme underlies catalysis, *Nature 438*, 117-121.

45. Schlegel, J., Armstrong, G. S., Redzic, J. S., Zhang, F., and Eisenmesser, E. Z. (2009) Characterizing and controlling the inherent dynamics of cyclophilin-A, *Protein Sci 18*, 811-824.

46. Agarwal, P. K., Geist, A., and Gorin, A. (2004) Protein dynamics and enzymatic catalysis: investigating the peptidyl-prolyl cis-trans isomerization activity of cyclophilin A, *Biochemistry 43*, 10605-10618.

47. Hammes-Schiffer, S., and Benkovic, S. J. (2006) Relating Protein Motion to Catalysis, *Annual Review of Biochemistry 75*, 519-541.

48. Bhabha, G., Lee, J., Ekiert, D. C., Gam, J., Wilson, I. A., Dyson, H. J., Benkovic, S. J., and Wright, P. E. (2011) A dynamic knockout reveals that conformational fluctuations influence the chemical step of enzyme catalysis, *Science 332*, 234-238.

49. Labeikovsky, W., Eisenmesser, E. Z., Bosco, D. A., and Kern, D. (2007) Structure and dynamics of pin1 during catalysis by NMR, *J Mol Biol 367*, 1370-1381.

50. Namanja, A. T., Peng, T., Zintsmaster, J. S., Elson, A. C., Shakour, M. G., and Peng, J. W. (2007) Substrate recognition reduces side-chain flexibility for conserved hydrophobic residues in human Pin1, *Structure 15*, 313-327.

51. Namanja, A. T., Wang, X. J., Xu, B., Mercedes-Camacho, A. Y., Wilson, K. A., Etzkorn, F. A., and Peng, J. W. (2011) Stereospecific gating of functional motions in Pin1, *Proc Natl Acad Sci U S A 108*, 12289-12294.

52. Pisliakov, A. V., Cao, J., Kamerlin, S. C., and Warshel, A. (2009) Enzyme millisecond conformational dynamics do not catalyze the chemical step, *Proc Natl Acad Sci U S A 106*, 17359-17364.

53. Kamerlin, S. C., and Warshel, A. (2010) At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis?, *Proteins 78*, 1339-1375.

54. Case, D. A., Cheatham, T. E., 3rd, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *Journal of computational chemistry 26*, 1668-1688.

55. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins 65*, 712-725.

56. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1996) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 5179−5197, *Journal of the American Chemical Society 118*, 2309-2309.

57. Doshi, U., and Hamelberg, D. (2009) Reoptimization of the AMBER Force Field Parameters for Peptide Bond (Omega) Torsions Using Accelerated Molecular Dynamics, *The Journal of Physical Chemistry B 113*, 16590-16595.

58. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *Journal of Chemical Physics 79*, 926-935.

59. Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes, *J. Comput. Phys. 23*, 327 - 341.

60. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) A smooth particle mesh Ewald method, *Journal of Chemical Physics 103*, 8577-8593.

61. Homeyer, N., Horn, A., Lanig, H., and Sticht, H. (2006) AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine, *Journal of Molecular Modeling 12*, 281-289.

62. Shen, T., and Hamelberg, D. (2008) A statistical analysis of the precision of reweighting-based simulations, *Journal of Chemical Physics 129*, 034103.

2.8 Figures

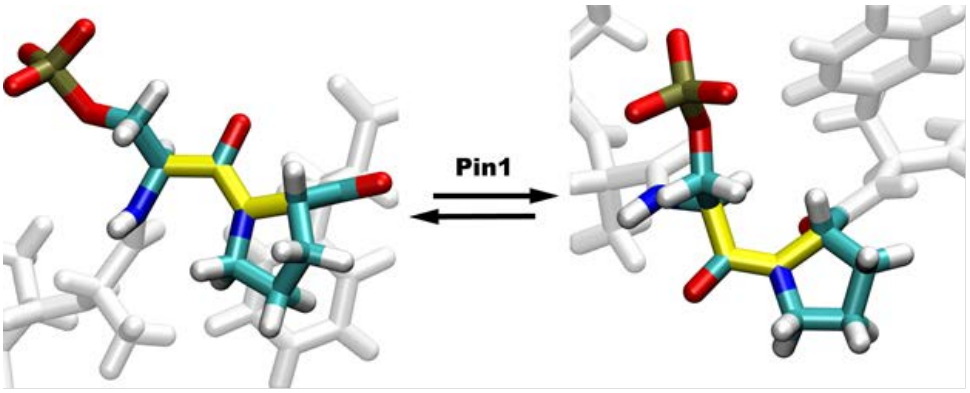

Figure 1. Pin1 catalyzes cis-trans isomerization of the ω-bond angle of proline when the preceding residue is a phosphorylated serine or phosphorylated threonine. The ω-bond angle is shown in yellow when it is in the trans (left) and cis (right) configuration

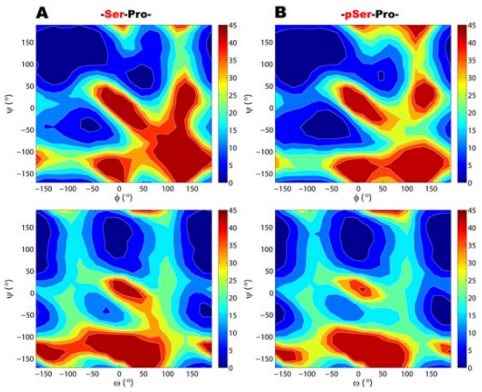

Figure 2. Contour plot in kcal/mol of the conformational space of unphosphorylated serine and phosphorylated serine in the free substrate, Ace-Ala-Ala-Ser/pSer-Pro-Phe-Nme. (A) The Ramachandran plot (top) of Ser and ω-ψ plot (bottom).

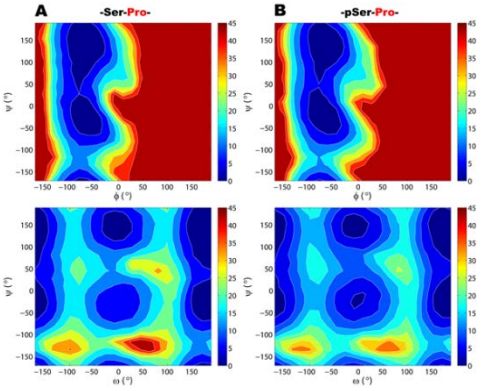

Figure 3. Contour plot in kcal/mol of the conformational space of Pro in the free phosphorylated and unphosphorylated substrate analogue. (A) The φ-ψ (top) and ω-ψ plot (bottom) of proline and the preceding ω-bond angle of the pSer-Pro motif

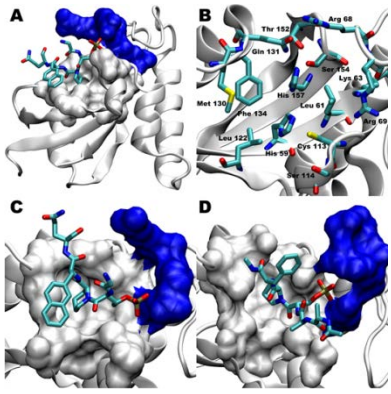

Figure 4. Structure and binding site of the catalytic domain of Pin1. (A) X-ray crystal structure (PDB ID 2Q5A) of the catalytic domain of Pin1 (shown as white), with a peptidomimetic in the active site. The active residues are shown using surface representation. . (B) The residues that make up the active sites. (C) The peptidomimetic inhibitor is shown interacting with the proline binding pocket and the phosphate binding residues (Lys 63, Arg 68, and Arg 69), shown as

blue. (D) The substrate analogue, Ace-Ala-Ala-pSer-Pro-Phe-Nme, is shown in the active site interacting with the proline binding pocket and the phosphate binding residues(Lys 63, Arg 68, and Arg 69), shown as blue.

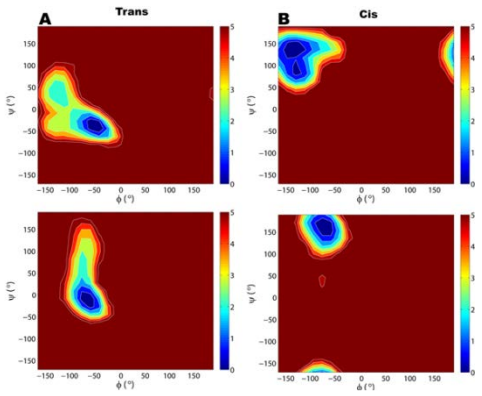

Figure 5. Contour plot in kcal/mol of the φ-ψ space or Ramachandran plot of the pSer-Pro motif of the substrate analogue, Ace-Ala-Ala-pSer-Pro-Phe-Nme, in the active site of the catalytic domain of Pin1 for pSer (top) and Pro (below), (A) when the ω-bond angle of the pSer-Pro motif is in the trans configuration and (B) when the ω-bond angle of the pSer-Pro motif is in the cis configuration.

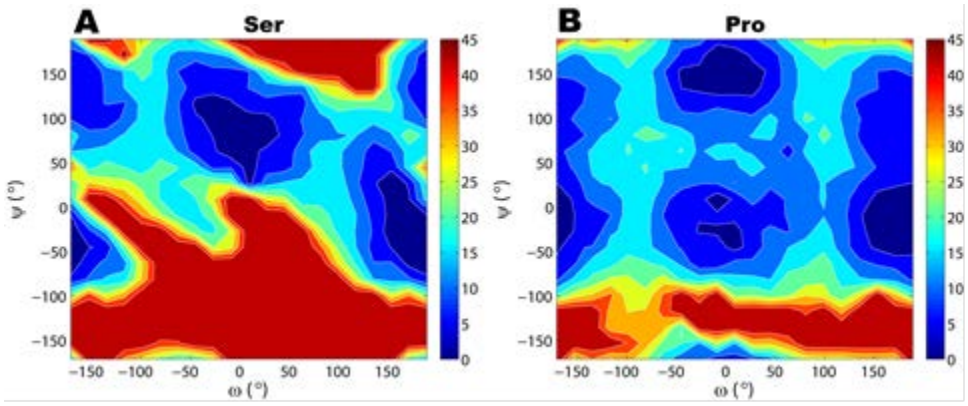

Figure 6. Contour plot in kcal/mol of the ω-ψ space of (A) pSer and (B) Pro of the pSer-Pro motif of the substrate analogue, Ace-Ala-Ala-pSer-Pro-Phe-Nme, when bound in the active site of the catalytic domain of Pin1.

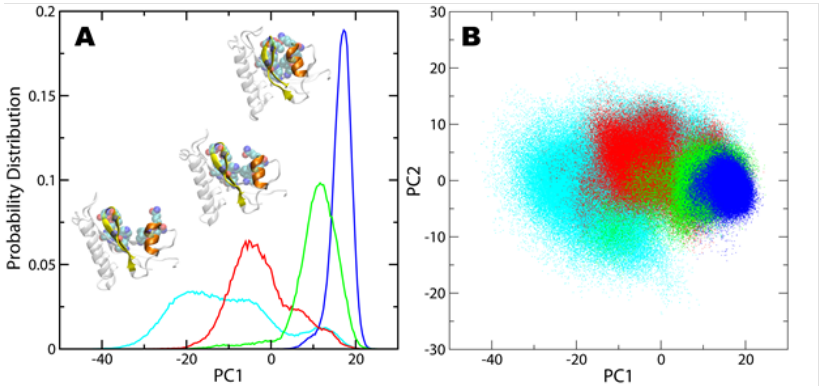

Figure 7. (A) Probability distribution of the first principal component (PC1) of the Pin1-substrate complex when the ω-bond angle of the pSer-Pro motif of the substrate is in the trans (blue), cis (green), and transition states (red) configurations and free Pin1

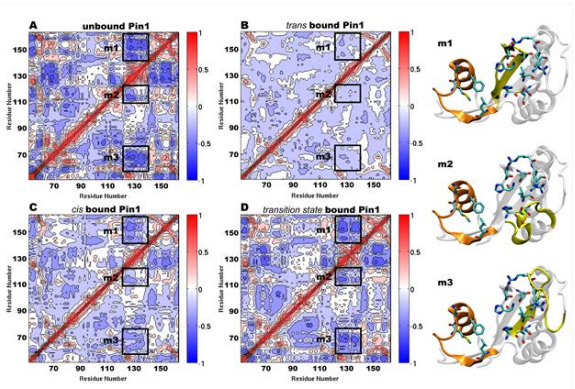

Figure 8. Dynamical cross correlated motions of Pin1 residues in the (A) free Pin1, and when bound to the substrate in the (B) trans, (C), cis, and (D) transition state configurations. The main correlated motions of Pin1 are labeled as m1, m2, and m3, and the re residues and domains involved in these motions are depicted on the right.

3 CONFORMATION DIRECTED CATALYSIS AND COUPLED ENZYME-SUBSTRATE DYNAMICS IN PIN1 PHOSPHORLATION DEPENDENT CIS-TRANS ISOMERASE

3.1 Abstract

Human peptidyl-prolyl *cis-trans* isomerase NIMA-interacting 1 (Pin1) is an essential enzyme in numerous phosphorylation dependent regulatory pathways and has been implicated in many diseases, including cancer and Alzheimer's. Pin1 specifically catalyzes the *cis-trans* isomerization of prolyl-peptide bonds preceded by phosphorylated serine or phosphorylated threonine in its protein substrates. Yet, little is known about the catalytic mechanism of Pin1 in atomistic detail. Here, we present results from accelerated molecular dynamics simulations to show that catalysis occurs along a restricted path of the backbone configuration of the substrate, selecting out specific conformations of the substrate in the active site of Pin1. We show that the dynamics of Pin1 and the enzyme-substrate interactions are intricately coupled to isomerization during catalysis. The strength of the interactions between the phosphate-binding pocket of Pin1 and the phosphate moiety of the substrate is dictated by the state of the substrate during catalysis. We also show that the transition state configuration of the substrate binds better to the catalytic domain of Pin1 relative to the *cis* and *trans* states, suggesting that Pin1 catalyzes its substrate by non-covalently stabilizing the transition state. These results suggest an atomistic detail understanding of the catalytic mechanism of Pin1 that is necessary to the design of novel inhibitors and the treatment of several diseases.

3.2 Introduction

The study of molecular switches on the atomistic level is important to the understanding of numerous biological processes. In general, biomolecular switches are regulated by many factors in order to ensure proper spatial and temporal functions, including post-translational modification, allosteric regulation, and transient interactions with cellular enzymatic components.[1-5] Pin1 (Figure 1) is one such cellular component that specifically recognizes phosphorylated serine/threonine (pSer/pThr)-proline motifs in its protein substrates and catalyzes the notoriously slow inter-conversion of the peptidyl-prolyl bond (ω-bond) from *cis* to *trans* and vice versa, a switching mechanism that is associated with many cellular phosphorylated dependent regulatory processes.[6-7] Pin1 therefore maintains the equilibrium between the *cis* and *trans* configurations of its phosphorylated substrates, in order to ensure proper function of other cellular components. For example, Pin1 ensures proper function of proline directed kinases and phosphatases that preferentially phosphorylate and dephosphorylate, respectively, serine/threonine-proline motifs when the peptide ω-bond is in a specific configuration ($\omega \sim \pm 180^{o}$ or 0^{o}).[8-10]

The specificity for phosphorylated serine/threonine-proline motifs makes Pin1 intricately involved in mitosis.[6, 11] Pin1 plays a role in oncogenesis[7, 12] and is overexpressed in the early stages of various cancers.[8, 13] Pin1 has also been implicated in the onset of Alzheimer's disease. It has been shown to regulate hyper-phosphorylated tau, a microtubule binding protein that is linked to plaque formations in brains of Alzheimer's patients.[14-16] Deregulation of Pin1 in neuronal cells has therefore been shown to initiate tau tangles.[8, 16] Interestingly, Pin1 has also been implicated in viral replication of the hepatitis C virus.[17] Pin1 has attracted tremendous attention as a target for the development of anticancer drugs[18-26]; therefore, an atomistic level

understanding of the catalytic mechanism of Pin1 is critical in designing novel inhibitors and the treatment of several diseases.

Pin1 is a two domain protein, containing a catalytic domain and a WW domain.[27] The presence of the WW domain does not greatly influence catalytic activity or binding of substrates to the catalytic domain.[28-29] Although Pin1 has been implicated in various diseases, its catalytic mechanism on the atomic level is still not well understood. Catalysis of *cis-trans* isomerization by Pin1 has been suggested, however, to proceed without any bond formation or breakage, similar to other peptidyl-prolyl *cis-trans* isomerases (PPIases).[30-31] X-ray crystallographic and NMR experiments suggest that preferential binding of the transition state is the primary factor driving Pin1 catalysis.[29, 32] Experimentally, it has been shown that the free energy barrier for *cis-trans* isomerization of a free substrate in solution is approximately 20.3 kcal/mol (84.8 kJ/mol), separating the cis and the *trans* states.[29, 33] The free energy barrier of the catalyzed *cis* to *trans* isomerization by Pin1 has been estimated to be approximately 13.2 kcal/mol (55.6 kJ/mol),[29] resulting in approximately 7.1 kcal/mol preferential binding of the transition state over the cis state of the substrate. A recent QM/MM study also suggested a similar catalytic free energy barrier.[34]

These studies suggest that the interactions between Pin1 and its protein substrates are crucial for catalysis, since these interactions dictate the affinity of the transition state relative to the *cis* and *trans* ground state configurations of the substrate. Several earlier studies have identified hydrogen bonding networks, hydrophobic interactions, and electrostatic interactions that are involved in maintaining the fold of the active site and in recognizing the protein substrates in Pin1 and related PPIases.[35-38] The phosphate-binding site is mainly made up of Lys 63, Arg 68, and Arg 69. Daum et al.[28] have shown that mutation of Arg68 to Ala does not affect

binding and catalytic activity as much as mutation of Arg 69 to Ala. Also, Behrsin et al.[31] have shown that Lys63 is critical for function. They showed that a dramatic loss of function was observed for a Lys63Ala substitution as compared to mutating Arg 68 or Arg 69. How are these interactions changing during catalysis? How does Pin1 take advantage of these interactions in order to preferentially stabilize the transition state?

In an attempt to shed light on these questions and provide an atomistic description of the catalytic mechanism of Pin1, we have carried out extensive accelerated molecular dynamics (aMD)[39] simulations on the substrate-bound Pin1 complex. Accelerated molecular dynamics allows us to sample the *cis-trans* interconversion within the active site of Pin1 and measure key atomistic features of the complex that are not accessible to current experimental techniques, including the short-lived transition state of the enzyme-substrate complex. Additionally, we have carried out normal molecular dynamics simulations on key sub-states of the enzyme-substrate complex in order to estimate the relative binding free energy using the molecular mechanics Poisson Boltzmann/Surface Area approach (MM/PBSA).[40] These computational studies provide a deeper insight into the mechanism of Pin1, and PPIases in general, furthering our understanding of the dynamical coupling of the interactions between enzymes and their substrates during catalysis.

3.3 Computational Methods

All of the simulations were carried out using the AMBER10 software suite[41] and the ff99SB[42] modified version of the Cornell et al. force field,[43] along with the re-optimized dihedral parameters for the peptide bond angle.[44] The force field parameters for the dianionic form of phosphorylated serine by Homeyer et al. were used.[45] The dianionic form of the phosphoserine residue is the preferred protonation state of the phosphate group of the substrate when bound to

Pin1.[33] The substrate analogue of Pin1 used in this study was Ace-Ala-Ala-pSer-Pro-Phe-Nme. The initial coordinates were obtained from a 1.5Å resolution X-ray crystal structure of a Pin1-inhibitor complex with PDB ID 2Q5A.[46] The initial structure of the Pin1-substrate complex was prepared as previously described[37] by modifying the peptidomimetic inhibitor to the substrate analogue. The WW domain was removed and only the catalytic domain was used in the simulations[28]. The two histidines in the active site were modeled in their mono-protonation state, as also suggested by x-ray crystallography.[35] The system was solvated with the TIP3P[47] water model in a periodic octahedron box. The SHAKE algorithm[48] was used to constrain all bonds involving hydrogen. The Langevin thermostat with a collision frequency of 1 ps^{-1} was used to maintain the temperature of the system at 300K during the simulations.[49] The electrostatic interactions were calculated using the particle mesh Ewald Summation method[50] and a cutoff of 9 Å was used for the direct non-bonded interactions. A time step of 2 fs was used to integrate Newton's equation of motion. All of the simulations were carried out using the NPT ensemble at a constant pressure of 1 bar. The system was equilibrated at 300 K and 1 bar, as previously described.[37]

Accelerated molecular dynamics (aMD)[39] simulations were carried out in order to observe *cis-trans* isomerization of the peptidyl prolyl bond angle in the substrate in the active site of Pin1 and free in solution, since the timescale of isomerization is extremely slow and beyond the timescale of normal molecular dynamics. A modified version of the pmemd module in Amber 10 was used to carry out all of the aMD simulations. The boost potential E was set to 127 kcal/mol and α was set to 15 kcal/mol. The potential energy of the total dihedral of the substrate was boosted during the aMD.[39] Snapshots during the simulations were saved every 5 steps in order to improve the statistics of the reweighted free energy profile, as previously

discussed.[51-52] Each configuration carried a weight of $e^{\beta\Delta V(r)}$, where $\Delta V(r)$ is the difference between the modified and unmodified potentials for that particular configuration. The probability distributions obtained from the aMD simulations were therefore reweighted using $e^{\beta\Delta V(r)}$, in order to calculate the distribution on the unmodified potential, $p(\xi)$, along one or two degrees of freedom, ξ. The free energy profiles were estimated using $-RT\ln[p(\xi)]$. The probability distributions of all of the 2D free energy profiles were calculated using a bin size of 20°x20°. Each bin was incremented by the weight of the configuration (frame), $e^{\beta\Delta V(r)}$, whenever ξ of the configuration (frame) fell inside the bin. In normal MD, which is equivalent to $\Delta V(\mathbf{r}) = 0$ for all $\mathbf{r}$, one would normally increment each bin by 1 whenever the ξ of the configuration falls within the confines of the bin. The 2D free energy profiles were plotted using the MATLAB program. The plots were normalized such that regions with the lowest free energy (most sampled regions) were assigned 0 kcal/mol and colored deep blue, and the free energies of the rest of the plot were relative to the lowest free energy regions, colored from blue to red. Un-sampled regions were also assigned the highest free energy and colored deep red. Five independent simulations of the Pin1-substrate complex were carried, each for 260 ns, for a total of 1.3 μs of simulation time. Also, ten independent simulations of the free substrate in solution were carried out, each for 260 ns, for a total of 2.6 μs of simulation time.

3.4 Results and Discussion

3.4.1 Conformation Directed Catalysis by Pin1

Enzymes are known to provide an environment where biochemical reactions can take place at a biologically relevant timescale, several orders of magnitude faster than the free spontaneous reaction in solution.[53] Generally, each family of enzymes is specific for a particular type of reaction, but a common theme, wherein enzymes provide a lower free energy reaction

path as the dominant contribution to catalysis, has emerged over the years.[54] In addition, the inherent dynamical nature of the enzyme is important and necessary, allowing for conformational rearrangements of the enzyme as the substrate goes through the transition state to the product state.[38, 55-59] Previously, we showed that Pin1 recognizes it substrates through conformational selection and binds specific sub-populations of the unbound substrate in the *cis*(ω ~ 0^o), transition state(ω ~ $+90^o$), and *trans*(ω ~ $\pm 180^o$) configurations.[37] Here, we have investigated the catalytic mechanism of Pin1 and focused on the changes in the free energy landscape of the substrate in the active site of Pin1 as compared to that free in solution, from extensive accelerated molecular dynamics simulations.

The Pin1 substrate analogue (Ace-Ala-Ala-pSer-Pro-Phe-Nme) can populate a number of conformational states free in solution, as projected on the reaction coordinate of the chemical step, the ω-bond angle, and the ψ-backbone dihedral angle of proline or the ψ-backbone dihedral angle of the phosphorylated serine (Figure 2A, B). The resulting 2D free energy landscape defined by the two backbone dihedral angles is excellent for describing prolyl *cis*-*trans* inter-conversion.[37, 60] In the active site of Pin1, the conformational space of the substrate is restricted and the catalytic domain of Pin1 shows intricate requirements for the *cis* and *trans* configurations (Figure 2C, D). Isomerization in the active site of Pin1 strictly proceeds from cis-α to trans-α, and vice versa, defined by ω-angle and the ψ-backbone dihedral angle of proline (Figure 2C). The backbone conformation of proline can go from cis-β to cis-α, and vice versa, more easily, while the population of the trans-β is tremendously reduced in the active site (Figure 2C). Similarly, but in a somewhat more restrictive manner, the backbone conformation of the preceding phosphorylated serine undergoes a concerted transition from cis-β to trans-α, and vice versa (Figure 2D). The cis-α of pSer is hardly populated free in solution and in the active site of

Pin1, while the population of the trans-β of pSer is dramatically reduced in the active site of Pin1.

The 1D free energy profile along the prolyl ω-bond angle is also shown in Figure 3A. The free energy barrier for the *cis-trans* isomerization in the active site of Pin1 is approximately 13 kcal/mol, compared to approximately 20 kcal/mol for the uncatalyzed process (Figure 3). These results therefore suggest that Pin1 can stabilize the transition state by as much as 7 kcal/mol, in excellent agreement to experiment[29, 33] and previous QM/MM estimates.[34] Similarly, it is seen that isomerization during catalysis in the active site of Pin1 is asymmetric and prefers to go through the $\omega \sim +90^{o}$ transition state (Figure 3A), as was observed for another PPIase (cyclophilin A).[61-62] The free energy barriers for the uncatalyzed and catalyzed isomerization are similar to the estimates from Figure 2. In addition to the conformational restricted path that is defined by Pin1 during catalysis, the results show that Pin1 binds the substrate in the transition state ($\omega \sim +90^{o}$) better than the *cis* configuration, which binds better than the *trans* configuration (Figure 3B). These results therefore suggest that the catalytic mechanism of Pin1 is mainly due to the preferential binding of the substrate in the transition state and non-covalently catalyzing the *cis-trans* isomerization of the prolyl ω-bond angle of phosphorylated Ser/Thr-Pro motifs.

3.4.2 Interactions between Pin1 and the phosphate moiety are key to stabilizing the transition state

What interactions are primarily responsible for stabilizing the transition state of the substrate? Figures 2 and 3 suggest that Pin1 binds the substrate in the transition state better than the *cis* and *trans* ground states of the substrate. Posttranslational phosphorylation of the -Ser-Pro- motif within the protein substrate renders the phosphorylated substrate specific for Pin1 recognition and catalysis.[63] Investigating the role of the phosphate group of the substrate and its

dynamics with the phosphate-binding pocket is therefore important to fully understanding the mechanism of Pin1. Here, we investigate the interactions between the phosphate moiety during catalysis and the phosphate-binding pocket of Pin1 that is made up of Lys63, Arg68 and Arg69 (Figure 4A) and to some extent Ser154 (Figure 5). Interestingly, our results generally suggest that the strength of the electrostatic interactions between the active site of phosphate binding site of Pin1 and the phosphate group of the substrate is coupled to the configuration of the substrate.

Unlike Arg69 and Lys63, Arg68 does not show any preference for any particular state of the substrate (Figure 4B). Regardless of the configuration of the substrate, the distance between the phosphorus atom of the phosphate group of the substrate and the carbon atom of the guanidinium moiety of Arg68 ranges from 3 to 17 Å, without any particular preference (Figure 4B). Figure 4B shows that Arg68 can equally form short- and long-ranged interactions with the substrate. Earlier mutagenesis studies[28, 31] have suggested that Arg 68 is not as crucial for substrate recognition, in complete agreement with our atomistic description of the dynamics of this residue in the complex.

On the other hand, Arg69 and Lys63 form tighter interactions with the phosphate group of the substrate in all of the configurations. More importantly, Lys 63 and Arg 69 interact differently with the different configurations of the substrate, as shown in Figure 4C, D. When the substrate is in the *cis* configuration, the distance between Arg 69 and the phosphate group ranges from approximately 3 to 13 Å. When the substrate is in the *trans* configuration the interaction distance ranges from 3 to 7 Å. However, when the substrate is in the transition state configuration, the intermolecular interaction distance is in a narrow 3 to 5 Å range. Similar interaction distances are observed between the nitrogen of the side chain of Lys63 and the phosphorus atom of the phosphate group of the substrate (Figure 4D). Our results therefore

suggest that the intermolecular interactions between the substrate and Lys63 and Arg69 are sensitive to the conformational state of the substrate. The enzyme forms more favorable electrostatic interactions with the phosphate group of the transition state configuration of the substrate than that of the cis and trans states. These preferential interactions with the transition state configuration by Pin1 are partly responsible for the lower free energy of binding the transition state.

We also identify a specific interaction between the side chain of Ser154 and the phosphate group of the substrate, as shown in Figure 5. This interaction is formed only when the substrate is in the *cis* and transition state configurations, but not when the substrate is in the *trans* configuration. The interaction of Ser154 with the phosphate group of the substrate when the substrate is in the transition state is tighter than that when the substrate is in the *cis* state. This interaction is therefore also coupled and sensitive to the state of the substrate and helps to stabilize the transition state, but it is expected to contribute much less than the interactions between Arg69 and Lys63 and the phosphate group. When the substrate is in the *trans* configuration, the hydroxyl group on the side chain of Ser154 instead forms a hydrogen bond with the carbonyl oxygen atom of the prolyl ω-bond angle of the substrate. Our results therefore suggest that the Ser154 is also important in stabilizing the transition state of the substrate through this configuration-dependent interaction with the phosphate group of the substrate.

Our results suggest that the electrostatic interactions between Pin1 and the phosphate group of the substrate are mainly responsible for stabilizing the transition state relative to the *cis* and *trans* states of the substrate. In addition, our results suggest that these interactions are coupled to the dynamics of the substrate and are strongest when the substrate is in the transition state. It has been shown experimentally that mutating Arg69 and Lys63 can severely compromise

the catalytic efficiency of Pin1.[28, 31] Therefore, exploiting the ability of Pin1 to discriminate between the different configurations of the substrate can aid in inhibitor design through the creation of transitions state analogues or the use of surrogate phosphate groups that can mimic the same types of interactions observed in the transition state.

3.4.3 Several intermolecular Pin1-substrate and intramolecular Pin1 interactions are coupled to the dynamics of the substrate

In addition to the observed coupling between the interactions in the phosphate binding pocket and the substrate, we show that other key interactions that may be important for substrate recognition are also sensitive to the state of the substrate. Cys113 is a key residue in the active site of Pin1 that is believed to make important contact with the substrate. The role of Cys113 in the active site of Pin1 is not well understood and has been a source of controversy. The controversy is partly due to the fact that -SH hydrogen bonds are generally not well understood[64] and the sulfur atom of the unprotonated form of the side chain of Cys can potentially act as a nucleophile. Therefore, Ranganathan et al. originally suggested that Cys113 is a nucleophile in the catalytic mechanism of Pin1, attacking the carbonyl carbon of the peptidyl-prolyl bond and altering the pseudo double bond character of the C-N bond to a single bond for a faster rotation in the active site of Pin1.[27] However, a unigenic evolution study has called this nucleophilic attack hypothesis into question.[31] Experimental studies have shown that mutation of Cys113 to Ser compromises but does not abolish isomerase activity of Pin1, suggesting that this residue plays an important role in Pin1 catalysis.[31, 36] We show that Cys113 can form hydrogen bonds with both the carbonyl oxygen of the prolyl peptide bond of the substrate and His59 of Pin1, switching between the two (Figure 6A, B). Cys113 can form a hydrogen bond with the oxygen atom of the prolyl-peptide bond when the substrate is in the *cis* state (Figure 6A). This hydrogen

bond persists in the transition state in a narrow range (ω ~ 90°, Figure 6A). The hydrogen bond however is never formed when the substrate is in the *trans* state (Figure 6A). Furthermore, Cys113 can also form an intramolecular hydrogen bond with His59 (Figure 6B). Our results indicate that the hydrogen bond is more localized when the substrate is in the transition state (Figure 6B). Our results therefore suggest that Cys113 is involved in substrate recognition and maintaining the fold of the active site. Also, His59 and His157 have been shown to be important for maintaining the fold and the shape of the active site of Pin1.[36] Our results therefore provide an atomistic description of these experimental findings, suggesting that Cys113 is also involved in maintaining the shape of the binding pocket through the hydrogen bond with His 59 and recognizing the substrate.

Interestingly, along with the enzyme-substrate intermolecular interactions, several of the intramolecular interactions in the enzyme are also coupled to the dynamics of the substrate, as is shown with the interaction between Cys113 and His59. Figure 6C shows that the side chain amine group of Gln131 prefers to interact with the carbonyl oxygen of the peptide bond when the substrate is in the *trans* configuration. Also, the enzyme forms an intramolecular interaction between Thr152 and Gln131. These two residues are parts of two different folds on either side of the proline-binding pocket and are coupled to the chemical step (ω-bond) of the substrate. The hydrogen bond between the hydroxyl group of Thr152 and oxygen atom of the side chain carbonyl of Gln131 is more stable when the substrate in the transition state than the *cis* and *trans* ground states, as shown in Figure 6D. In general, the interactions are more localized when the substrate is in the transition state, forming a well-defined and compact active site to recognize the transition state. Our present results suggest that intermolecular Pin1-substrate interactions and intramolecular Pin1 hydrogen bonds are sensitive and coupled to the dynamics of the

substrate and serve multiple roles, including creating necessary interactions for recognizing the different states of substrate, maintaining the structural integrity of the binding pocket, and orchestrating the overall dynamics of the enzyme that is necessary for catalysis.

3.4.4 Pin1 preferentially binds the transition state configuration of its substrate

The above results suggest that Pin1 binds its substrate when it is in the transition state more tightly than the *trans* and *cis* states, as evident in Figures 2 and 3. This preferential binding of the transition state is mainly due to the localized interactions that are observed when the substrate is in the transition state. In order to further investigate the ability of Pin1 to preferentially bind the transition state, we have carried out three 50ns normal molecular dynamics simulations of the Pin1-substrate complex when the substrate is in the *cis*-α, transition state and *trans*-α configurations. The substrate was maintained in the transition state by keeping the ω-bond angle of the prolyl-peptide bond ~90° using a flat bottom potential with a force constant of 1000 kcal/mol/rad^2. During the 50 ns normal MD simulations of the substrate in the *cis*-α and *trans*-α wells, the ψ-angle of the prolyl residue of the substrate was held between -50° and 50° with the same force constant. The trajectories were analyzed using molecular mechanics/Poisson Boltzmann Surface Area (MM/PBSA) method,[40] in order to directly estimate the binding free energies of the different configurations of the substrates in the active site of Pin1. The changes in configuration entropy were not taken into consideration in these estimates, and the relative changes were assumed to be negligible. We show that Pin1 binds the transitions state more favorably than the *cis* and *trans* states (Figure 7). Figure 7 shows the distributions of the free energies of binding the different states of the substrate by Pin1. Preferentially binding the transition state by Pin1 over the *cis* and *trans* states thus results in a lower free energy barrier of *cis-trans* isomerization in the active site. These results also suggest that the *cis*-α state of the

substrate binds to Pin1 slightly better than the *trans*-α state of the substrate, inline with previous experimental results showing that the *cis* configuration of the substrate binds better than the *trans* configuration.[20, 29, 58] Energy decomposition analysis shows that both the polar and non-polar interactions are most favorable for transition state stabilization. However, binding of the transition state configuration results in the highest electrostatic desolvation penalty, as one would expect, since the phosphate moiety of the transition state configuration of the substrate forms very tight electrostatic interactions with the phosphate-binding pocket of Pin1.

3.5 Conclusions

We have carried out extensive accelerated molecular dynamics simulations and free energy calculations in order to investigate the catalytic mechanism of Pin1. Pin1 selectively binds sub-microstates of its substrate, resulting in a substrate conformation-directed catalytic mechanism. The dynamics of the Pin1 is intricately coupled to the dynamics of the substrate during catalysis. We show enzyme-substrate interactions are extremely sensitive to the configuration of the substrate, preferentially forming tighter interactions with the transition state configuration. Pin1 therefore binds the transition state configuration of the substrate more favorable than the *cis* and *trans* configurations, resulting in a lower free energy barrier during catalysis compared to the spontaneous isomerization in solution. Our all-atom simulation results are consistent with previous experiments and provide an atomistic description of the catalytic mechanism of Pin1, details that complement current experimental techniques. The simulations provide a level of detail that could aid in the development of inhibitors, possibly taking advantage of the fact that the enzyme can discriminate between subsets of ensemble of substrate configurations.

3.6 Acknowledgement

This work is supported in part by the National Science Foundation CAREER (MCB-0953061), Georgia Research Alliance, and Georgia State University. This work was also supported by Georgia State's IBM System p7 supercomputer, acquired through a partnership of the Southeastern Universities Research Association and IBM supporting the SURAgrid initiative.

3.7 References

1. Deribe, Y. L.; Pawson, T.; Dikic, I., Post-translational modifications in signal integration. *Nat. Struct. Mol. Biol.* **2010,** *17*, 666-672.

2. Goodey, N. M.; Benkovic, S. J., Allosteric regulation and catalysis emerge via a common route. *Nat. Chem. Biol.* **2008,** *4*, 474-482.

3. Kuriyan, J.; Eisenberg, D., The origin of protein interactions and allostery in colocalization. *Nature* **2007,** *450*, 983-990.

4. Ha, J.-H.; Loh, S. N., Protein Conformational Switches: From Nature to Design. *Chemistry – A European Journal* **2012,** *18*, 7984-7999.

5. Tsai, C. J.; del Sol, A.; Nussinov, R., Allostery: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol.* **2008,** *378*, 1-11.

6. Kun Ping, L.; Finn, G.; Tae Ho, L.; Nicholson, L. K., Prolyl cis-trans isomerization as a molecular timer. *Nat. Chem. Biol.* **2007,** *3*, 619-629.

7. Theuerkorn, M.; Fischer, G.; Schiene-Fischer, C., Prolyl cis/trans isomerase signalling pathways in cancer. *Curr. Opin. Pharmacol.* **2011,** *11*, 281-287.

8. Lu, K. P.; Zhou, X. Z., The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nat. Rev. Mol. Cell Biol.* **2007,** *8*, 904-916.

9. Zhou, X. Z., et al., Pin1-dependent prolyl isomerization regulates dephosphorylation of Cdc25C and tau proteins. *Mol. cell* **2000,** *6*, 873-83.

10. Werner-Allen, J. W.; Lee, C. J.; Liu, P.; Nicely, N. I.; Wang, S.; Greenleaf, A. L.; Zhou, P., cis-Proline-mediated Ser(P)5 dephosphorylation by the RNA polymerase II C-terminal domain phosphatase Ssu72. *J. Biol. Chem.* **2011,** *286*, 5717-26.

11. Lu, K.; Hanes, S.; Hunter, T., A human peptidyl-prolyl isomerase essential for regulation of mitosis. *Nature* **1996,** *380*, 544-7.

12. Ryo, A.; Liou, Y.-C.; Lu, K. P.; Wulf, G., Prolyl isomerase Pin1: a catalyst for oncogenesis and a potential therapeutic target in cancer. *J. Cell Sci.* **2003,** *116*, 773-783.

13. Finn, G.; Lu, K. P., Phosphorylation-specific prolyl isomerase Pin1 as a new diagnostic and therapeutic target for cancer. *Curr. Canc. Drug Targ.* **2008,** *8*, 223-229.

14. Bulbarelli, A.; Lonati, E.; Cazzaniga, E.; Gregori, M.; Masserini, M., Pin1 affects Tau phosphorylation in response to A beta oligomers. *Mol. Cell. Neurosci.* **2009,** *42*, 75-80.

15. Pei-Jung, L.; Wulf, G.; Zhou, X.; Davies, P.; Lu, K., The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* **1999,** *399*, 784-8.

16. Pastorino, L., et al., The prolyl isomerase Pin1 regulates amyloid precursor protein processing and amyloid-beta production. *Nature* **2006,** *440*, 528-34.

17. Lim, Y.-S.; Tran, H. T. L.; Park, S.-J.; Yim, S.-A.; Hwang, S. B., Peptidyl-Prolyl Isomerase Pin1 Is a Cellular Factor Required for Hepatitis C Virus Propagation. *J. Virol.* **2011,** *85*, 8777-8788.

18. Mori, T., et al., A dual inhibitor against prolyl isomerase Pin1 and cyclophilin discovered by a novel real-time fluorescence detection method. *Biochem. Biophys. Res. Commun.* **2011,** *406*, 439-443.

19. Potter, A. J., et al., Structure-guided design of α-amino acid-derived Pin1 inhibitors. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 586-590.

20. Wang, X. D. J.; Xu, B. L.; Mullins, A. B.; Neiler, F. K.; Etzkorn, F. A., Conformationally locked isostere of phosphoSer-cis-Pro inhibits Pin1 23-fold better than phosphoSer-trans-Pro isostere. *J. Am. Chem. Soc.* **2004,** *126*, 15533-15542.

21. Liu, T.; Liu, Y.; Kao, H.-Y.; Pei, D., Membrane Permeable Cyclic Peptidyl Inhibitors against Human Peptidylprolyl Isomerase Pin1. *J. Med. Chem.* **2010,** *53*, 2494-2501.

22. Wildemann, D.; Erdmann, F.; Alvarez, B. H.; Stoller, G.; Zhou, X. Z.; Fanghänel, J.; Schutkowski, M.; Lu, K. P.; Fischer, G., Nanomolar Inhibitors of the Peptidyl Prolyl Cis/Trans Isomerase Pin1 from Combinatorial Peptide Libraries. *J. Med. Chem.* **2006,** *49*, 3430-3430.

23. Potter, A., et al., Discovery of cell-active phenyl-imidazole Pin1 inhibitors by structure-guided fragment evolution. *Bioorg. Med. Chem. Lett.* **2010,** *20*, 6483-6488.

24. Lv, L.; Zhou, Z.; Huang, X.; Zhao, Y.; Zhang, L.; Shi, Y.; Sun, M.; Zhang, J., Inhibition of peptidyl–prolyl cis/trans isomerase Pin1 induces cell cycle arrest and apoptosis in vascular smooth muscle cells. *Apoptosis* **2010,** *15*, 41-54.

25. Hennig, L.; Christner, C.; Kipping, M.; Schelbert, B.; Rücknagel, K. P.; Grabley, S.; Küllertz, G.; Fischer, G., Selective Inactivation of Parvulin-Like Peptidyl-Prolyl cis/trans Isomerases by Juglone. *Biochemistry* **1998,** *37*, 5953-5960.

26. Zhu, L.; Jin, J.; Liu, C.; Zhang, C.; Sun, Y.; Guo, Y.; Fu, D.; Chen, X.; Xu, B., Synthesis and biological evaluation of novel quinazoline-derived human Pin1 inhibitors. *Biorg. Med. Chem.* **2011,** *19*, 2797-2807.

27. Ranganathan, R.; Lu, K. P.; Hunter, T.; Noel, J. P., Structural and Functional Analysis of the Mitotic Rotamase Pin1 Suggests Substrate Recognition Is Phosphorylation Dependent. *Cell* **1997,** *89*, 875-886.

28. Daum, S.; Fanghänel, J.; Wildemann, D.; Schiene-Fischer, C., Thermodynamics of Phosphopeptide Binding to the Human Peptidyl Prolyl cis/trans Isomerase Pin1. *Biochemistry* **2006,** *45*, 12125-12135.

29. Greenwood, A. I.; Rogals, M. J.; De, S.; Lu, K. P.; Kovrigin, E. L.; Nicholson, L. K., Complete determination of the Pin1 catalytic domain thermodynamic cycle by NMR lineshape analysis. *J. Biomol. NMR* **2011,** *51*, 21-34.

30. Fanghanel, J.; Fischer, G., Insights into the catalytic mechanism of peptidyl prolyl cis/trans isomerases. *Front. Biosci.* **2004,** *9*, 3453-3478.

31. Behrsin, C. D.; Bailey, M. L.; Bateman, K. S.; Hamilton, K. S.; Wahl, L. M.; Brandl, C. J.; Shilton, B. H.; Litchfield, D. W., Functionally Important Residues in the Peptidyl-prolyl Isomerase Pin1 Revealed by Unigenic Evolution. *J. Mol. Biol.* **2007,** *365*, 1143-1162.

32. Xu, G. Y. G.; Zhang, Y.; Mercedes-Camacho, A. Y.; Etzkorn, F. A., A Reduced-Amide Inhibitor of Pin1 Binds in a Conformation Resembling a Twisted-Amide Transition State. *Biochemistry* **2011,** *50*, 9545-9550.

33. Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J.-U.; Lu, K. P.; Fischer, G., Role of Phosphorylation in Determining the Backbone Dynamics of

the Serine/Threonine-Proline Motif and Pin1 Substrate Recognition. *Biochemistry* **1998,** *37*, 5566-5575.

34. Vöhringer-Martinez, E.; Duarte, F.; Toro-Labbe, A., How Does Pin1 Catalyze the Cis-Trans Prolyl Peptide Bond Isomerization? A QM/MM and Mean Reaction Force Study. *J. Phys. Chem. B* **2012,** *116*, 12972-12979.

35. Mueller, J. W.; Link, N. M.; Matena, A.; Hoppstock, L.; Ruppel, A.; Bayer, P.; Blankenfeldt, W., Crystallographic Proof for an Extended Hydrogen-Bonding Network in Small Prolyl Isomerases. *J. Am. Chem. Soc.* **2011,** *133*, 20096-20099.

36. Bailey, M. L.; Shilton, B. H.; Brandl, C. J.; Litchfield, D. W., The Dual Histidine Motif in the Active Site of Pin1 Has a Structural Rather than Catalytic Role. *Biochemistry* **2008,** *47*, 11481-11489.

37. Velazquez, H. A.; Hamelberg, D., Conformational Selection in the Recognition of Phosphorylated Substrates by the Catalytic Domain of Human Pin1. *Biochemistry* **2011,** *50*, 9605-9615.

38. McGowan, Lauren C.; Hamelberg, D., Conformational Plasticity of an Enzyme during Catalysis: Intricate Coupling between Cyclophilin A Dynamics and Substrate Turnover. *Biophys. J.* **2013,** *104*, 216-226.

39. Hamelberg, D.; Mongan, J.; McCammon, J. A., Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004,** *120*, 11919-11929.

40. Kollman, P. A., et al., Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000,** *33*, 889-897.

41. Case, D. A., et al., The Amber biomolecular simulation programs. *J. Comput. Chem.* **2005,** *26*, 1668-1688.

42. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* **2006,** *65*, 712-725.

43. Cornell, W. D., et al., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 5179−5197. *J. Am. Chem. Soc.* **1996,** *118*, 2309-2309.

44. Doshi, U.; Hamelberg, D., Reoptimization of the AMBER Force Field Parameters for Peptide Bond (Omega) Torsions Using Accelerated Molecular Dynamics. *J. Phys. Chem. B* **2009,** *113*, 16590-16595.

45. Homeyer, N.; Horn, A.; Lanig, H.; Sticht, H., AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *J. Mol. Model.* **2006,** *12*, 281-289.

46. Zhang, Y., et al., Structural Basis for High-Affinity Peptide Inhibition of Human Pin1. *ACS Chem. Biol.* **2007,** *2*, 320-328.

47. Jorgensen, W. L., Revised TIPS for simulations of liquid water and aqueous solutions. *J. Chem. Phys.* **1982,** *77*, 4156-4163.

48. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977,** *23*, 327 - 341.

49. Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D., Langevin stabilization of molecular dynamics. *J. Chem. Phys.* **2001,** *114*.

50. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995,** *103*, 8577-8593.

51. Shen, T.; Hamelberg, D., A statistical analysis of the precision of reweighting-based simulations. *J. Chem. Phys.* **2008,** *129*, 034103-9.

52. Doshi, U.; Hamelberg, D., Improved Statistical Sampling and Accuracy with Accelerated Molecular Dynamics on Rotatable Torsions. *J Chem Theory Comput* **2012,** *8*, 4004-4012.

53. Wolfenden, R.; Snider, M. J., The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res.* **2001,** *34*, 938-945.

54. Wolfenden, R., Thermodynamic and extrathermodynamic requirements of enzyme catalysis. *Biophys. Chem.* **2003,** *105*, 559-572.

55. Wolfenden, R., Enzyme catalysis: Conflicting requirements of substrate access and transition state affinity. *Mol. Cell. Biochem.* **1974,** *3*, 207-211.

56. Ma, B.; Nussinov, R., Enzyme dynamics point to stepwise conformational selection in catalysis. *Curr. Opin. Chem. Biol.* **2010,** *14*, 652-659.

57. Fraser, J. S.; Clarkson, M. W.; Degnan, S. C.; Erion, R.; Kern, D.; Alber, T., Hidden alternative structures of proline isomerase essential for catalysis. *Nature* **2009,** *462*, 669-673.

58. Namanja, A. T.; Wang, X. J.; Xu, B.; Mercedes-Camacho, A. Y.; Wilson, K. A.; Etzkorn, F. A.; Peng, J. W., Stereospecific gating of functional motions in Pin1. *Proc. Nat. Acad. Sci.* **2011,** *108*, 12289-12294.

59. Doshi, U.; McGowan, L. C.; Ladani, S. T.; Hamelberg, D., Resolving the complex role of enzyme conformational dynamics in catalytic function. *Proc. Nat. Acad. Sci.* **2012,** *109*, 5699-5704.

60. Fischer, S.; Dunbrack, R. L.; Karplus, M., Cis-Trans Imide Isomerization of the Proline Dipeptide. *J. Am. Chem. Soc.* **1994,** *116*, 11931-11937.

61. Hamelberg, D.; McCammon, J. A., Mechanistic Insight into the Role of Transition-State Stabilization in Cyclophilin A. *J. Am. Chem. Soc.* **2008,** *131*, 147-152.

62. Ladani, S. T.; Hamelberg, D., Entropic and Surprisingly Small Intramolecular Polarization Effects in the Mechanism of Cyclophilin A. *J. Phys. Chem. B* **2012,** *116*, 10771-10778.

63. Yaffe, M. B., et al., Sequence-specific and phosphorylation-dependent proline isomerization: a potential mitotic regulatory mechanism. *Science* **1997,** *278*, 1957-1960.

64. Zhou, P.; Tian, F.; Lv, F.; Shang, Z., Geometric characteristics of hydrogen bonds involving sulfur atoms in proteins. *Proteins: Struct., Funct., Bioinf.* **2009,** *76*, 151-163.

3.8 Figures

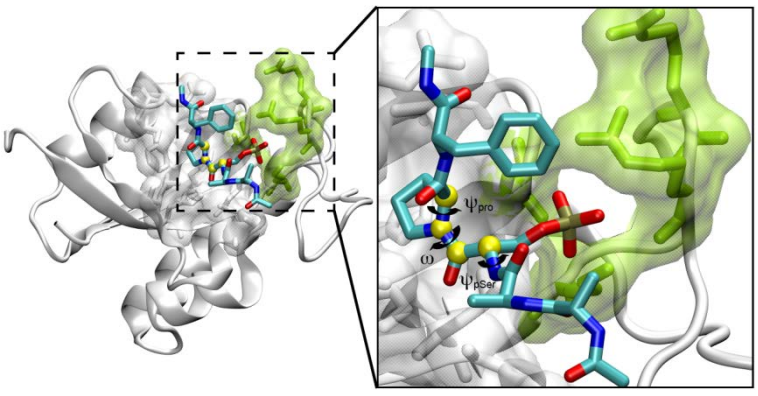

Figure 9. Pin1 and the substrate analogue (AApSPF) in the active site. A close-up of the active site is shown and the relevant backbone dihedral angles of the substrate are also defined. The phosphate-binding site is colored using green transparent surface representation and the residues are shown using the green stick representation. The residues shown

in green are Lys 63, Arg 68, Arg 69, and Ser 154. The rest of the active site is shown using a white transparent surface representation and stick model.

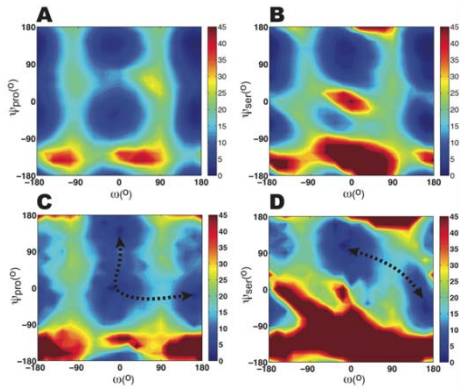

Figure 10. The phase space of the -pSer-Pro- motif in the substrate analogue free in solution and in the active site of Pin1. Free energy landscapes in kcal/mol along ω-ψ of Pro (A) and pSer (B) free in solution and Pro (C) and pSer (D) in the active site of Pin1. Free energy landscapes in kcal/mol along ω-ψ of Pro (A) and pSer (B) free in solution and Pro (C) and pSer (D) in the active site of Pin1. The dashed lines with arrows in (C) and (D) depict the lowest free energy paths along the backbone of Pro and pSer during catalysis.

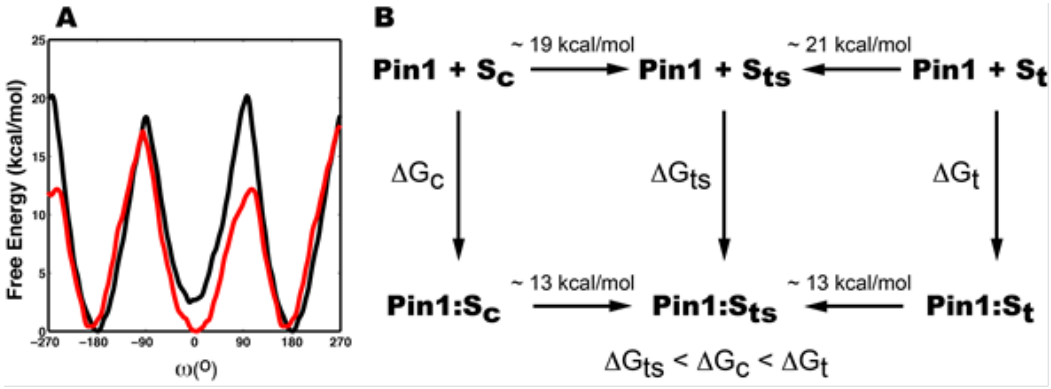

Figure 11. Free energy of cis-trans isomerization of the substrate in solution and the active site of Pin1. (A) Free energy profiles in kcal/mol along the ω-bond angle of the –pSer-Pro- motif in the substrate analogue for the uncatalyzed isomerization reaction (b lack) and the catalyzed isomerization reaction by Pin1(red). (B) Thermodynamic cycle connecting the isomerization process in solution and active site of Pin1 with the free energies of binding the cis (c), transition (ts) and trans (t) states of the substrate.

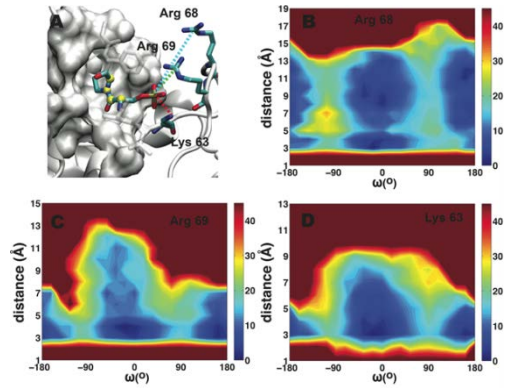

Figure 12. Dynamical coupling between the phosphate group of the substrate and Pin1 phosphate-binding site during catalysis. The residues in the phosphate binding site of Pin1 and the interactions made with the substrate (A). 2D free energy profiles in kcal/mol along the ω-bond of the -pSer-Pro- motif of the substrate and the distance between the guanidinium carbon of Arg 68 and the phosphorus atom of the phosphate group (B), the distance between the guanidinium carbon of Arg 69 and the phosphorus atom of the phosphate group (C), and the distance between the side chain nitrogen atom of Lys 63 and the phosphorus atom of the phosphate group (C).

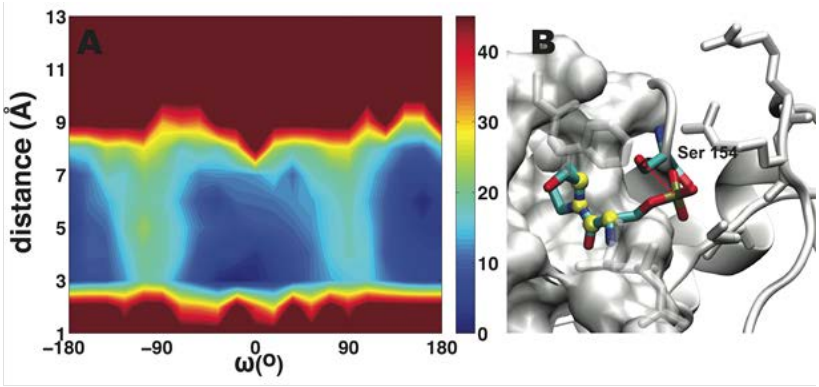

Figure 13. Dynamical coupling between the phosphate group of the substrate and Ser154, in the active site of Pin1, during catalysis. 2D free energy profiles in kcal/mol along the ω-bond of the -pSer-Pro- motif of the substrate and the distance between the side ch oxygen atom of Ser 154 and the phosphorus atom of the phosphate group (A). Ser 154 in the active site of Pin1 and the substrate analogue in the active site are shown (B).

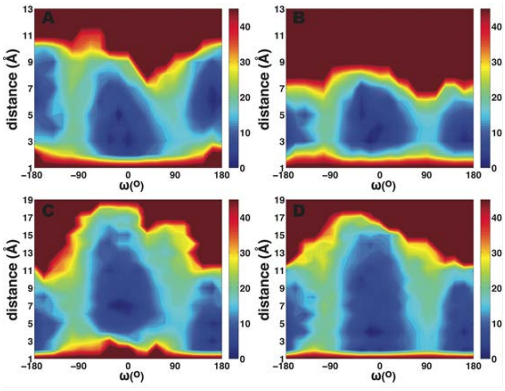

Figure 14. Dynamical coupling between the substrate and Pin1 during catalysis. 2D free energy profiles in kcal/mol along the ω-bond of the -pSer-Pro- motif of the substrate and the distance between the sulfur atom of Cys113 and the carbonyl oxygen atom of the pe ptide ω-bond of the substrate (A), the distance between the sulfur atom of Cys113 and the ε-nitrogen of His59 (B), the distance between the side chain nitrogen atom of Gln131 and the carbonyl oxygen atom of the peptide ω-bond of the substrate (C), and the distance between the side chain oxygen atom of Gln131 and the side chain oxygen atom of Thr152 (D).

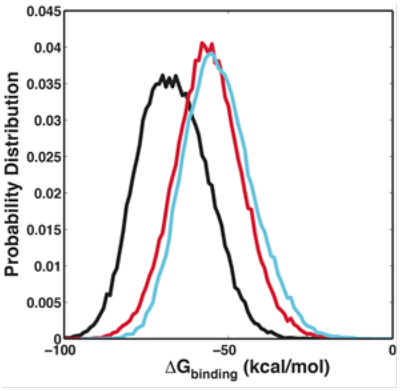

Figure 15. Probability distributions of free energies of binding the substrate in different configurations. Free energies in kcal/mol of binding the transitions state (black), cis-α (red) and trans-α (blue) configurations in the active site of Pin1.

4 KINETIC RATES AND DYNAMICAL EFFECTS OF PHOSPHORYLATION ON PEPTIDYL PROLYL CIS-TRANS ISOMERIZATION FROM CONSTANT FORCE MOLECULAR DYNAMICS

4.1 Abstract

The cis-trans isomerization of proline plays an important role in various biological processes. It is a rate limiting step in the protein folding process because it is slow on the biological timescale. Atomic Force Microscopy (AFM) experiments show that prolyl cis-trans isomerization can be catalyzed by a mechanical force such that the "mechanically catalyzed" kinetic rate makes prolyl

isomerization a tractable molecular switch. In order to fully understand “mechanically catalyzed” prolyl isomerization, properly extracting kinetic information from experimental or computational data is pivotal. Two popular models have emerged to extract kinetic information from AFM experiments. The original model proposed by Bell is later extended by Dudko, Hummer, and Szabo. In this study, we use constant force biased molecular dynamics simulations to evaluate the popular Bell model and the extension to it proposed by Dudko, Hummer, and Szabo. We extract kinetic information using both models for the X-Pro motif where X is either a phosphorylated or unphosphorylated serine residue. The results show that phosphorylation retards isomerization of the X-Pro motif. Interestingly, we observed that phosphorylation does not change the free energy barrier of the X-Pro motif. The autocorrelation functions were calculated for both the X=Ser and X=pSer polypeptides to shed light on this observation. The autocorrelation functions of the X=Ser and X=pSer polypeptides show that the phosphorylated polypeptide diffuses more slowly than the unphosphorylated analogue resulting in the slower kinetic rate for the phosphorylated peptide reconciling the calculated kinetic rate constants with the free energy barriers.

4.2 Introduction

Proline is a unique member of the 20 common amino acids. Its side chain is tied into its backbone creating conformational restrictions that do not exist in the other amino acids[1]. The cis configuration of the peptidyl-prolyl peptide bond is nearly isoenergetic with the trans state[2]. The pseudo double bond character of the peptide bond causes the barrier for cis-trans isomerization to be ~17-20 kcal/mol[2], resulting in notoriously slow kinetics[2].The slow kinetics of prolyl cis-trans isomerziation are a rate-limiting step in protein folding[3]. Nature uses the resulting structural differences between a protein with a cis peptidyl prolyl bond and that with a trans peptidyl prolyl

bond to regulate biological functions. A class of structurally unrelated family of enzymes, generally known as peptidyl-prolyl isomerases (PPIases), catalyzes the cis-trans isomerization of the peptidyl- proline bond in order to increase the rate of isomerization to a biologically relevant timescale[4].

Additionally, peptidyl-prolyl cis-trans isomerization can be coupled to post-translational phosphorylation in order to facilitate more complex forms of molecular recognition and regulation in biological systems[5]. Phosphorylation alters molecular recognition based on the perturbation the phosphate group creates on the conformational spaces and the electrostatic environments of polypeptides relative to the unphosphorylated case[6,7,8,9,10]. Pin1, a phosphorylation dependent PPIase, specifically recognizes phosphorylated serine/threonine motifs in its protein substrates[11]. This stringent specificity and the synergy between cis-trans isomerization and posttranslational phosphorylation make Pin1 important in the regulation of the proline directed kinases and phosphotase[12,13,14], cancer[15], Alzheimer's[16], and many other diseases[12]. The functional role of the peptidyl-prolyl bond as a molecular switch has rekindled a great deal of interest in fully understanding the mechanism of peptidyl-prolyl cis-trans isomerization and PPIases in biological processes.

The majority of known proteins that undergo prolyl cis-trans isomerization are regulated by PPIases. However, some proteins undergo uncatalyzed peptidyl prolyl isomerization at rates comparable to those the catalyzed reaction. Ribonuclease A, for example, can fold in ~70 ms without a PPIase, even though the folding depends on cis-trans isomerization of Pro93, Pro114, and Pro117[17]. The giant muscle protein titin can also achieve this feat[18,19]. It can do this because it needs to have a mechanical force act upon it in order to perform its function. These results suggest that internal molecular forces could speed up the rate of isomerization. Using Atomic

Force Microscopy, it was shown that a mechanical force can speed up the rate of prolyl cis-trans isomerization in the VPGXG repeats in elastin like proteins[20]. Molecular dynamics simulations have also been used to study mechanically induced cis-trans isomerization in elastic proteins[21, 22]. These results suggest that peptidyl prolyl isomerization during the folding of ribonuleaseA[17] can be achieved via “mechanical catalysis”. These findings provide a possible molecular mechanism for muscle contractions in proteins such as titin[20]. This ability to undergo cis-trans isomerization under tension underscores the fact that proteins can also experience tension under crowded physiological conditions[23]. AFM experiments have therefore been valuable in understanding folding-unfolding and other properties of biomolecules[24] under tension and have been complemented by computational simulations[25].

Previously, extracting meaningful kinetic information from AFM experiments and complementary computational simulations was challenging. A popular method for extracting kinetic information from experimental and computational studies was the Bell's formalism[26,27,28]. More recently, however, an extension to Bell's model was proposed by Dudko, Hummer, and Szabo[29,30,31] which takes into account fluctuations in the reaction coordinate caused by the force applied to the system by the cantilever. In the present study, we probe the cis to trans transition of the peptidyl-prolyl bond of the phosphorylated serine-proline motif in a well-studied peptide substrate of Pin1 under constant biasing force molecular dynamics simulations. The applied constant pulling force accelerates the rate of the cis to trans transition of prolyl isomerization that is otherwise too slow to be observed using normal molecular dynamics simulations. We carried out simulations of the unphosphorylated and phosphorylated substrates using several constant-pulling forces. This allows us to understand the effect of phosphorylation on the “mechanical catalysis” of the peptidyl-prolyl bond. The Dudko-Hummer-Szabo (DHS) model is used to fit the

data and probe the rates of isomerization of the phosphorylated and unphosphorylated substrate under no force. The results are then compared to experiments. It also allows us to test the predictability of kinetics using classical molecular mechanics with dihedral parameters that were optimized from a thermodynamic point of view. The data is also fitted to the more simplistic Bell model for comparison. This study provides a detailed atomistic description of the effect of phosphorylation of the kinetics of peptidyl-prolyl isomerization that is in agreement with experiments and has implications for molecular switches and the mechanism of action of Pin1.

4.3 Computational Methods

All simulations were carried out using the Amber 10 suite of programs [32] and the modified parm99SB[33] version of the Cornell et al. force field[34], with the re-optimized dihedral parameters for the peptide backbone ω-bond[35]. The force field parameters for the dianionic phosphorylated serine (pSer) used in this study were taken from the work by Homeyer et al.[36] The NPT ensemble at a pressure of 1 bar and a temperature of 300K was used for all the simulations with explicit TIP4P-EW[37] or TIP3P[38] water models, and the SHAKE[39] algorithm was used to constrain all bonds involving hydrogen. The Langevin thermostat[40] was used to regulate the temperature of the system at 300 K with a collision frequency of 1.0 ps-1. Short-range nonbonded interactions were calculated with a cutoff of 9 Å, and all long range interactions were calculated using the particle mesh Ewald summation[41]. A timestep of 2 fs was used to integrate the Langevin equation of motion.

4.3.1 Accelerated molecular dynamics of prolyl peptide substrates

The unphosphorylated and phosphorylated substrates, Ace-Ala-Ala-X-Pro-Phe-Nme, were built using the XLEAP module in the Amber software package. The substrates were placed in a cubic periodic TIP3P[38] water box with the edges of the box at least 10 Å away from any part of the

substrates. Two Na^+ ions were added to the systems with the phosphorylated substrate in order to attain electrostatic neutrality. The systems were then equilibrated with a series of minimization and molecular dynamics simulations, as previously described[42].

Accelerated molecular dynamics (aMD)[43] simulations were carried out in order to observe *cis-trans* isomerization of the peptidyl prolyl bond angle in the substrate in free solution because the timescale of isomerization is extremely slow and beyond the timescale of normal molecular dynamics for the purposes of characterizing the free energy landscape. An in house modified version of the pmemd module in Amber 10 was used to carry out all of the aMD simulations. The boost potential E was set to 127 kcal/mol and α was set to 15 kcal/mol. The total dihedral potential energy of the substrate was boosted during the aMD[43]. Trajectories were saved every 5 molecular dynamics steps. This is done to minimize the errors of the reweighted free energy profile, as previously discussed[44, 45].

Each snapshot of the trajectories carries a weight of $e^{\beta\Delta V(r)}$, where $\Delta V(r)$ is the difference between the modified and unmodified potentials for that specific snapshot. In order to calculate the distribution on the unmodified potential, $p(\xi)$, the probability distributions obtained from the aMD simulations were reweighted using $e^{\beta\Delta V(r)}$. After the trajectories were reweighted, the free energy profiles were calculated using $-RT\ln[p(\xi)]$. The probability distributions of the 2D free energy profiles were calculated using a bin size of 20°x20°. A bin size of 1° was used for the 1D free energy profile. Each bin was incremented by the weight of the configuration, $e^{\beta\Delta V(r)}$, whenever ξ of the configuration fell inside the bin. The free energy profiles were plotted using the MATLAB program. The plots were normalized such that regions with the lowest free energy were assigned 0 kcal/mol, and the free energies of the rest of the plot were relative to the lowest free energy regions. Un-sampled regions were also assigned the

highest free energy and colored deep red. Ten independent simulations of the substrate were carried out for each polypeptide substrate (X = Ser, pSer), each for 260 ns.

4.3.2 Simulations of the cis to trans transition under a constant pulling force

An ensemble of substrate conformations was created in order to generate a distribution of cis to trans transition times under a constant pulling force. The ensemble of starting structures was generated by a 5 ns normal molecular dynamics simulation of both the phosphorylated and unphosphorylated substrates in a periodic box of TIP4P-EW[37] water with the prolyl peptide bond (omega bond) in the cis state. The TIP4P-EW model was used in the kinetic studies because its diffusion coefficient is $2.4*10^{-5}$ cm^2/s[46]. This is close to the values obtained in experimental studies[47,48,49]. A snapshot of the conformation was saved every 500 molecular dynamics steps. Each substrate was simulated using five different constant forces: 60, 65, 70, 75, and 80 kcal/mol*Å. The force was applied on the alpha carbons of the proline and Ser/pSer residues as illustrated in Figure 1. At the beginning of each simulation for each force, a starting structure was randomly selected from the ensemble of structures generated from the 5 ns simulation of the cis state. The simulation was stopped as soon as the configuration of the peptidyl-prolyl bond transitioned to the trans configuration based on the C-alpha - C-alpha distance. The C-alpha - C-alpha distance for the trans state is ~3.5 Å. The survival time of the cis configuration was recorded. The C-alpha - C-alpha cutoff distance was determined from the results of the accelerated molecular dynamics simulations described in the previous section. Upon entering the trans well, the current simulation was stopped and a new simulation was started using another starting structure randomly selected from the ensemble of conformations using a random initial distribution of the velocities. The number of transitions observed from the cis to trans well for each constant pulling force is shown in Table 1.

4.3.3 Construction of the Survival Probability Distributions

The dwell time distributions of the cis isomer were recorded and counted. The dwell time is the time the peptide substrate spends in the cis configuration until it traverses the free energy barrier into the trans configuration. This results in a time series, T_i, where i= 1, 2, 3, …, N, where N is the number of times the substrate traverses the free energy barrier and escapes into the trans state. The survival function $S(t) = \int_t^\infty p(\tau)d\tau$ was calculated in order to obtain a distribution of dwell times. The probability that the system escapes the cis well at time T_i or longer was calculated as $S(T_i) \approx (i-1)/N$. These escape probabilities were then sorted and ordered from highest escape probability to smallest escape probability. All of the constant force biased simulations showed a similar exponential decay given by $S(t) = e^{-kt}$ where k is the decay constant and S(t) is the escape probability. Escape probability distributions have been previously shown to decay in this manner[50].

4.3.4 Fitting to the Dudko-Hummer-Szabo (DHS) Model

The variation of the survival time with the constant applied force can be fitted to the DHS model, as shown in equation 1[30], in order to estimate the survival time at zero force. Here τ represents the residence time of the cis state. The variable τ_0 represents the residence time of the cis state in the absence of a constant pulling force. The variable υ represents the shape of the energy barrier[40] and F represents the applied force. $X^{\ddagger}$ is the difference in the c-alpha – c-alpha distance when the substrate is in the cis basin from that when the substrate is at the transition state ($\Delta G^{\ddagger}$) along the free energy landscape defined by the c-alpha – c-alpha distance. The value of $x^{\ddagger}$ could be obtained from the fit equation 1. We determined $x^{\ddagger}$ from the free energy profile along the c-

alpha – c-alpha distance. After the survival times at each constant biasing force were determined, they were plotted for fitting to the DHS model. The fitted parameters are reported in Table 2.

$$\tau(F) = \tau_0 \left(1 - \frac{\upsilon F x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{1-\frac{1}{\upsilon}} e^{-\beta \Delta G^{\ddagger}\left[1-\left(1-\frac{1-\upsilon F x^{\ddagger}}{\Delta G^{\ddagger}}\right)^{\frac{1}{\upsilon}}\right]} \quad (1)$$

4.3.5 Calculation of the autocorrelation functions at 300K

In order to understand the physical basis for the differing kinetics between the phosphorylated and unphosphorylated substrates, the autocorrelation functions for the X = pSer and X = Ser were calculated at 300K. Since prolyl isomerization is an Ornstein-Uhlenbeck process[51], the autocorrelation function of ω is related to the diffusion constant via equation 2.

$$C(t) = \frac{\sum_{t=j+1}^{n}\left(y_t - \bar{y}\right)\left(y_{t-j} - \bar{y}\right)}{\sum_{t=1}^{n}\left(y_t - \bar{y}\right)^2} = \frac{e^{-\frac{tDK}{k_B T}}}{\frac{k_B}{K}} \quad (2)$$

4.4 Results and Discussion

4.4.1 The peptidyl-prolyl bond angle is correlated to the Cα-Cα distance

We carried accelerated molecular dynamics on the -Ser-Pro- containing peptide and observed cis-trans isomerization of the prolyl peptide bond. It has been prediviously shown that the proyl peptide bond (ω) angle itself is a good reaction coordinate for the cis-trans isomerization of proline (Figure 1)[7, 42,52,53]. We show in Figure 2 that the distance between the Cα of Ser and the Cα of Pro (Cα-Cα) is coupled to the ω-bond angle and also serves as a good reaction coordinate for prolyl cis-trans isomerization. When the peptide bond is in the trans configuration, the Cα-Cα distance is between 3.5 and 4.0 Å and when the peptide bond is in the cis configuration, the Cα-

Cα distance is mainly between 2.6 and 3.3 Å. The transition state, which is when the ω-bond is at ~90^{o}, is at a Cα-Cα distance of 3.3-3.4 Å. The free energy barrier separating the cis and trans configuration is ~17 kcal/mol, in good agreement with experimentally estimated free energy barriers for –Ser-Pro- motifs[54]. The peptide in the cis configuration therefore is slightly longer than the peptide in the trans configuration by ~ 0.8 Å per peptide bond. The distance from the cis configuration of a peptide bond to the transition state, $x^{\ddagger}$, along the Cα-Cα distance is ~ 0.3-0.4 Å. Proteins with several repeats of the -X-Pro- motif can therefore undergo elongation and contraction as they undergo cis-trans isomerization. The AFM experiments of Valiaev et al.[20] demonstrated the modulation in the length of an elastin like protein as it undergoes prolyl cis-trans isomerization.

4.4.2 Cis to trans isomerization under tension

Cis-trans isomerization cannot be observed during normal molecular dynamics because of the high barrier of ~17-20 kcal/mol separating the cis and the trans configurations. However, under constant pulling force along the Cα-Cα distance, cis to trans isomerization can be sped up as the constant external force effectively lowers the barrier between the cis and trans configurations. We carried out constant pulling force simulations at different pulling forces and monitored the change in the Cα-Cα distance as shown in Figure 3. This figure shows the raw data for one of the constant force biased simulations. The time of transition from cis to trans as the Cα-Cα distance crossed the transition state was recorded (Table 1). The simulation was repeated many times using a starting structure from the ensemble of cis configurations generated from the 5 ns normal molecular dynamics simulation of the cis state. The constant force simulations were carried out for the unphosphorylated and phosphorylated peptide using several constant forces listed in Table 1. These simulations allowed us to generate distributions of cis-trans isomerization times

under different pulling forces.

4.4.3 Extracting the survival time at zero force with the Dudko-Hummer-Szabo (DHS) model

The DHS model was used to extract kinetic information from the constant force biased simulations[30]. The survival times at each force are acquired from the integration of the survival probability distributions at each force displayed in Figure 4. This integration yields the residence time τ at that constant biasing force. The residence times at each force were then plotted for fitting to the DHS model shown in equation 1 (Figure 5). The DHS model has several variables which must be accounted for if the survival time in the absence of a biasing force is to be extracted. From Figure 2, we know that $x^{\ddagger}$ is ~0.4 Å. The value of this variable is taken to be the distance from the cis state to the transition state[29]. Therefore, we bounded $x^{\ddagger}$ in the fit to be between ~0.3 – 0.5 Å. This allows the value to be different for both the phosphorylated and unphosphorylated cases of the substrate. The fitted value for the substrate with the phosphorylated serine residue was 0.390 Å and the value for the substrate with the unphosphorylated serine residue was 0.383Å (Table 1). The free energy of the transition state ($\Delta G^{\ddagger}$) was bouded to the range of ~17-20 kcal/mol. This selection of this range of values is well supported by previous computational and experimental data[42,54,55]. Furthermore, this choice is also in agreement with the free energy surfaces constructed in Figure 2. The fitted transition state free energy for the phosphorylated substrate was 17.7 kcal/mol and 17.6 kcal/mol unphosphorylated substrate (Table 2). The fits yielded the survival times of 311.184 and 128.701 s for the X = pSer and X = Ser substrates, respectively (Table 2). These survival times correspond to rate constants of $3.214*10^{-3}$ and $7.76994*10^{-3}$ s^{-1} for the X = pSer and X = Ser substrates, respectively. The experimental values from Schutkowski et al. are $4.2*10^{-3}$ and

$9.7*10^{-3}$ s^{-1} for the phosphorylated and unphosphorylated polypeptides, respectivley[54]. From these results, it can be concluded that the addition of a phosphate group frustrates isomerization in agreement with experimental data[51]. Additionally, these results show that reliable kinetic information can be extracted directly from constant force biased molecular dynamics trajectories using the DHS model.

4.4.4 Extracting the survival time at zero force with the Bell Model

The relationship between mechanical forces and biological phenomena was first highlighted by Bell[26]. In order to extract the survival time of the cis state in the absence of a biasing force (τ_0), the survival time at each biasing force was plotted as it was for the DHS model (Figure 5). The data was fitted to the equation $\tau(F) = \tau_0 \exp(-Fx^{\ddagger})$ where τ_0 and $x^{\ddagger}$ are fitted parameters proposed by Bell[26]. The fit for parameter $x^{\ddagger}$ was bounded based on Figure 2 as it was for the fit to the DHS model. The results of the fitting yielded a survival time of 0.302191 s and 0.13081 s for the phosphorylated and unphosphorylated polypeptides, respectively. These survival times correspond to rate constants of 3.3092 and 7.6447 s^{-1} for the phosphorylated and unphosphorylated peptides, respectively. These values do not agree well with a previous experimental study of Pin1 substrates[54]. A previous QM/MM study noted that the Bell model may not be sufficient to extract kinetics from force biased trajectories[22].

4.4.5 Diffusion Accounts for the difference in the kinetic rate constants of the phosphorylated and unphosphorylated polypeptide substrates

It was interesting to observe that there is no apparent difference in the $\Delta G^{\ddagger}$ of the phosphorylated and unphosphorylated substrates (Figure 6). This led us to investigate where the difference in the rate constant between the phosphorylated and unphosphorylated substrates

arises. In order to tease out more kinetic information, Kramer's rate theory is invoked because it rigorously defines the Arrhenius pre-factor in terms of the diffusion coefficient and the curvature of the free energy surface[56]. Since the free energy plots of the polypeptide substrates are so similar, diffusion is the most likely place in Kramer's formalism for differences in the two substrates to appear. Therefore, the diffusion constant for each substrate is the main focus for reconciling the free energy plots of Figure 6 with the rate constants at zero force (τ_0) determined from the DHS fits (Figure 5).

It was previously shown that prolyl isomerization is an Ornstein-Uhlenbeck process[51,46]. Therefore, relationship between the autocorrelation function and the diffusion constant is described by equation 2. This means that the autocorrelation function can give insight about the diffusion constant of the omega bond for the phosphorylated and unphosphorylated substrates used in this study. Therefore, the autocorrelation for both the phosphorylated and unphosphorylated substrates were calculated at 300K from a 10 ns simulation of the trans state (Figure 7). Figure 7 shows that the phosphorylated substrate decays slower than the unphosphorylated substrate. The slower decay of the phosphorylated substrate is indicative of a slower diffusion constant via equation 2. This shows that the autocorrelation function reconciles the calculated kinetic constants with the free energy profiles of Figure 6.

4.5 Conclusions

This study has explored the kinetic properties of a Pin1 substrate and its unphosphorylated analogue through the use of constant force biased molecular dynamics. The Bell model and the extension to it originally proposed by Dudko, Hummer, and Szabo were both used to probe the kinetic properties of these two polypeptide substrates. The extension proposed by Dudko, Hummer, and Szabo resulted in superior agreement with experimental studies. An interesting

observation of this study was that the energy of the transition states of the polypeptide substrates were not substantially different despite the difference in their kinetic properties. To reconcile the one dimensional free energy surfaces in Figure 6 with the results of Table 2, the autocorrelation function of both polypeptide substrates were calculated because the autocorrelation function provides information about the diffusion constant. The autocorrelation functions indicate that the diffusion constant for the phosphorylated substrate is lower than that of the unphosphorylated substrate thereby reconciling the one dimensional free energy surface with kinetic data extracted from the DHS model.

4.6 References

1. Beck, D. A. C.; Alonso, D. O. V.; Inoyama, D.; Daggett, V., The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences* **2008,** *105* (34), 12259-12264.

2. Kun Ping, L.; Finn, G.; Tae Ho, L.; Nicholson, L. K., Prolyl cis-trans isomerization as a molecular timer. *Nature Chemical Biology* **2007,** *3* (10), 619-629.

3. Brandts, J. F.; Halvorson, H. R.; Brennan, M., Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry* **1975,** *14* (22), 4953-4963.

4. Göthel, S. F.; Marahiel, M. A., Peptidyl-prolyl cis-trans isomerases, a superfamily of ubiquitous folding catalysts. *Cellular and Molecular Life Sciences CMLS* **1999,** *55* (3), 423-436.

5. Zhou, X. Z.; Lu, P. J.; Wulf, G.; Lu, K. P., Phosphorylation-dependent prolyl isomerization: a novel signaling regulatory mechanism. *Cellular and Molecular Life Sciences* **1999,** *56* (9), 788-806.

6. Shen, T.; Wong, C. F.; McCammon, J. A., Atomistic Brownian Dynamics Simulation of Peptide Phosphorylation. *Journal of the American Chemical Society* **2001,** *123* (37), 9107-9111.

7. Hamelberg, D.; Shen, T.; McCammon, J. A., Phosphorylation Effects on cis/trans Isomerization and the Backbone Conformation of Serine−Proline Motifs: Accelerated Molecular Dynamics Analysis. *Journal of the American Chemical Society* **2005,** *127* (6), 1969-1974.

8. Tholey, A.; Lindemann, A.; Kinzel, V.; Reed, J., Direct Effects of Phosphorylation on the Preferred Backbone Conformation of Peptides: A Nuclear Magnetic Resonance Study. *Biophysical Journal* **1999,** *76* (1), 76-87.

9. Kyung-Koo, L.; Cheonik, J.; Seongeun, Y.; Hogyu, H.; Minhaeng, C., Phosphorylation effect on the GSSS peptide conformation in water: Infrared, vibrational circular dichroism, and circular dichroism experiments and comparisons with molecular dynamics simulations. *Journal of Chemical Physics* **2007,** *126* (23), 235102.

10. Byun, B. J.; Kang, Y. K., Conformational preferences and prolyl cis–trans isomerization of phosphorylated Ser/Thr-Pro motifs. *Biopolymers* **2010,** *93* (4), 330-339.

11. Adams, J. A., Kinetic and Catalytic Mechanisms of Protein Kinases. *Chemical Reviews* **2001,** *101* (8), 2271-2290.

12. Lu, K. P.; Zhou, X. Z., The prolyl isomerase PIN1: a pivotal new twist in phosphorylation signalling and disease. *Nature Reviews. Molecular Cell Biology* **2007,** *8* (11), 904-916.

13. Zhou, X. Z.; Kops, O.; Werner, A.; Lu, P.-J.; Shen, M.; Stoller, G.; Küllertz, G.; Stark, M.; Fischer, G.; Lu, K. P., Pin1-Dependent Prolyl Isomerization Regulates Dephosphorylation of Cdc25C and Tau Proteins. *Molecular Cell* **2000,** *6* (4), 873-883.

14. Werner-Allen, J. W.; Lee, C.-J.; Liu, P.; Nicely, N. I.; Wang, S.; Greenleaf, A. L.; Zhou, P., cis-Proline-mediated Ser(P)5 Dephosphorylation by the RNA Polymerase II C-terminal Domain Phosphatase Ssu72. *Journal of Biological Chemistry* **2011,** *286* (7), 5717-5726.

15. Theuerkorn, M.; Fischer, G.; Schiene-Fischer, C., Prolyl cis/trans isomerase signalling pathways in cancer. *Current Opinion in Pharmacology* **2011,** *11* (4), 281-287.

16. Pei-Jung, L.; Wulf, G.; Zhou, X.; Davies, P.; Lu, K., The prolyl isomerase Pin1 restores the function of Alzheimer-associated phosphorylated tau protein. *Nature* **1999,** *399*, 6738, 784-8.

17. Dodge, R. W.; Scheraga, H. A., Folding and Unfolding Kinetics of the Proline-to-Alanine Mutants of Bovine Pancreatic Ribonuclease A†. *Biochemistry* **1996,** *35* (5), 1548-1559.

18. Fowler, S. B.; Best, R. B.; Toca Herrera, J. L.; Rutherford, T. J.; Steward, A.; Paci, E.; Karplus, M.; Clarke, J., Mechanical Unfolding of a Titin Ig Domain: Structure of Unfolding Intermediate Revealed by Combining AFM, Molecular Dynamics Simulations, NMR and Protein Engineering. *Journal of Molecular Biology* **2002,** *322* (4), 841-849.

19. Maruyama, K., Connectin/titin, giant elastic protein of muscle. *The FASEB Journal* **1997,** *11* (5), 341-5.

20. Valiaev, A.; Lim, D. W.; Oas, T. G.; Chilkoti, A.; Zauscher, S., Force-Induced Prolyl Cis−Trans Isomerization in Elastin-like Polypeptides. *Journal of the American Chemical Society* **2007,** *129* (20), 6491-6497.

21. Shen, T.; Hamelberg, D.; McCammon, J. A., Elasticity of peptide omega bonds. *Physical Review E* **2006,** *73* (4), 041908.

22. Chen, J.; Edwards, S. A.; Gräter, F.; Baldauf, C., On the Cis to Trans Isomerization of

Prolyl–Peptide Bonds under Tension. *The Journal of Physical Chemistry B* **2012,** *116* (31), 9346-9351.

23. Christoph, A. H.; Rob, P., Directional interactions and cooperativity between mechanosensitive membrane proteins. *EPL (Europhysics Letters)* **2013,** *101* (6), 68002.

24. Clausen-Schaumann, H.; Seitz, M.; Krautbauer, R.; Gaub, H. E., Force spectroscopy with single bio-molecules. *Current Opinion in Chemical Biology* **2000,** *4* (5), 524-530.

25. Isralewitz, B.; Gao, M.; Schulten, K., Steered molecular dynamics and mechanical functions of proteins. *Current Opinion in Structural Biology* **2001,** *11* (2), 224-230.

26. Bell, G., Models for the specific adhesion of cells to cells. *Science* **1978,** *200* (4342), 618-627.

27. Evans, E.; Ritchie, K., Dynamic strength of molecular adhesion bonds. *Biophysical Journal* **1997,** *72* (4), 1541-1555.

28. Evans, E.; Berk, D.; Leung, A., Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments. *Biophysical Journal* **1991,** *59* (4), 838-848.

29. Dudko, O. K.; Hummer, G.; Szabo, A., Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Physical Review Letters* **2006,** *96* (10), 108101.

30. Dudko, O. K.; Hummer, G.; Szabo, A., Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proceedings of the National Academy of Sciences* **2008,** *105* (41), 15755-15760.

31. Hummer, G.; Szabo, A., Kinetics from nonequilibrium single-molecule pulling experiments. *Biophysical Journal* **2003,** *85* (1), 5-15.

32. Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J., The Amber biomolecular simulation programs. *Journal Of Computational Chemistry* **2005,** *26* (16), 1668-1688.

33. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C., Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics* **2006,** *65* (3), 712-725.

34. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules J. Am. Chem. Soc. 1995, 117, 5179−5197. *Journal of the American Chemical Society* **1996,** *118* (9), 2309-2309.

35. Doshi, U.; Hamelberg, D., Reoptimization of the AMBER Force Field Parameters for Peptide Bond (Omega) Torsions Using Accelerated Molecular Dynamics. *The Journal of Physical Chemistry B* **2009,** *113* (52), 16590-16595.

36. Homeyer, N.; Horn, A.; Lanig, H.; Sticht, H., AMBER force-field parameters for phosphorylated amino acids in different protonation states: phosphoserine, phosphothreonine, phosphotyrosine, and phosphohistidine. *Journal of Molecular Modeling* **2006,** *12* (3), 281-289.

37. Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T., Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics* **2004,** *120* (20), 9665-9678.

38. Jorgensen, W. L., Revised TIPS for simulations of liquid water and aqueous solutions. *The Journal of Chemical Physics* **1982,** *77* (8), 4156-4163.

39. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C., Numerical Integration of the Cartesian

Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977,** *23,* 327 - 341.

40. Izaguirre, J. A.; Catarello, D. P.; Wozniak, J. M.; Skeel, R. D., Langevin stabilization of molecular dynamics. *Journal of Chemical Physics* **2001,** *114* (5).

41. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G., A smooth particle mesh Ewald method. *The Journal of Chemical Physics* **1995,** *103* (19), 8577-8593.

42. Velazquez, H. A.; Hamelberg, D., Conformational Selection in the Recognition of Phosphorylated Substrates by the Catalytic Domain of Human Pin1. *Biochemistry* **2011,** *50* (44), 9605-9615.

43. Hamelberg, D.; Mongan, J.; McCammon, J. A., Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *Journal of Chemical Physics* **2004,** *120* (24), 11919-11929.

44. Shen, T.; Hamelberg, D., A statistical analysis of the precision of reweighting-based simulations. *The Journal of Chemical Physics* **2008,** *129* (3), 034103-9.

45. Doshi, U.; Hamelberg, D., Improved Statistical Sampling and Accuracy with Accelerated Molecular Dynamics on Rotatable Torsions. *Journal of Chemical Theory and Computation* **2012,** *8* (11), 4004-4012.

46. Johnson, Q.; Doshi, U.; Shen, T.; Hamelberg, D., Water's Contribution to the Energetic Roughness from Peptide Dynamics. *Journal of Chemical Theory and Computation* **2010,** *6* (9), 2591-2597.

47. Gillen, K. T.; Douglass, D. C.; Hoch, M. J. R., Self□Diffusion in Liquid Water to −31°C. *The Journal of Chemical Physics* **1972,** *57* (12), 5117-5119.

48. Holz, M.; Heil, S. R.; Sacco, A., Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate 1H NMR PFG measurements. *Physical Chemistry Chemical Physics* **2000,** *2* (20), 4740-4742.

49. Price, W. S.; Ide, H.; Arata, Y., Self-Diffusion of Supercooled Water to 238 K Using PGSE NMR Diffusion Measurements. *The Journal of Physical Chemistry A* **1999,** *103* (4), 448-450.

50. Hamelberg, D.; Shen, T.; Andrew McCammon, J., Relating kinetic rates and local energetic roughness by accelerated molecular-dynamics simulations. *The Journal of Chemical Physics* **2005,** *122* (24), -.

51. Wang, M. C.; Uhlenbeck, G. E., On the Theory of the Brownian Motion II. *Reviews of Modern Physics* **1945,** *17* (2-3), 323-342.

52. Yonezawa, Y.; Nakata, K.; Sakakura, K.; Takada, T.; Nakamura, H., Intra- and Intermolecular Interaction Inducing Pyramidalization on Both Sides of a Proline Dipeptide during Isomerization: An Ab Initio QM/MM Molecular Dynamics Simulation Study in Explicit Water. *Journal of the American Chemical Society* **2009,** *131* (12), 4535-4540.

53. Fischer, S.; Dunbrack, R. L.; Karplus, M., Cis-Trans Imide Isomerization of the Proline Dipeptide. *Journal of the American Chemical Society* **1994,** *116* (26), 11931-11937.

54. Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J.-U.; Lu, K. P.; Fischer, G., Role of Phosphorylation in Determining the Backbone Dynamics of the Serine/Threonine-Proline Motif and Pin1 Substrate Recognition†. *Biochemistry* **1998,** *37* (16),

5566-5575.

55. Greenwood, A. I.; Rogals, M. J.; De, S.; Lu, K. P.; Kovrigin, E. L.; Nicholson, L. K., Complete determination of the Pin1 catalytic domain thermodynamic cycle by NMR lineshape analysis. *Journal Of Biomolecular NMR* **2011,** *51* (1-2), 21-34.

56. Hänggi, P.; Talkner, P.; Borkovec, M., Reaction-rate theory: fifty years after Kramers. *Reviews of Modern Physics* **1990,** *62* (2), 251-341.

4.7 Figures

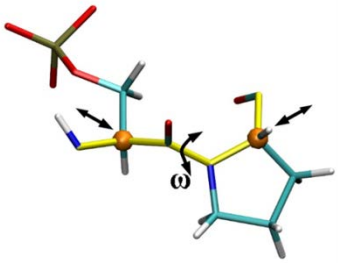

Figure 16. Prolyl cis-trans isomerization about the omega bond (ω). The alpha carbons of the X residue and the prolyl residue are represented as orange spheres. The biasing force for the constant force biased simulations were applied to the alpha carbons of the X residue and the prolyl residue are represented as orange spheres. The biasing force for the constant force biased simulations were applied to the alpha carbons of the X and prolyl residues as illustrated by the arrows.

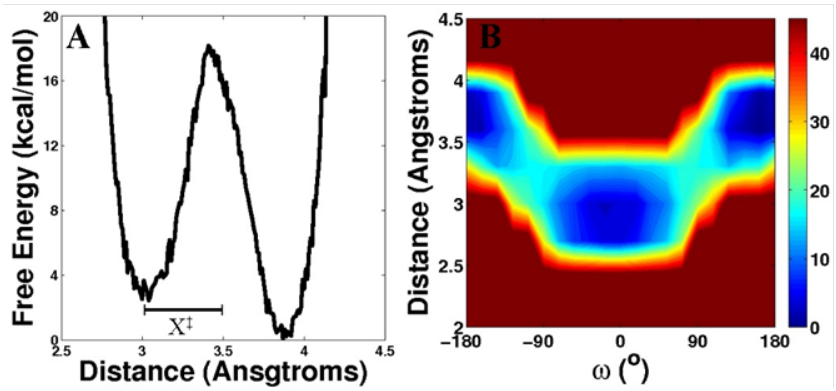

Figure 17. Free Energy Profiles showing the coupling of the omega bond to the c-alpha - c-alpha distance. Panel A shows how the energy fluctuates with the distance between the alpha carbons of the X and prolyl residues. Panel B shows the coupling between the omeg a bond and the distance between the alpha carbons of the X and prolyl residues.

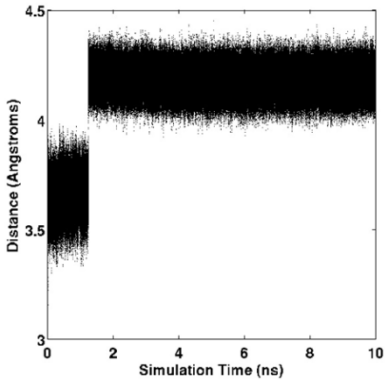

Figure 18. Constant force bias simulation of the X = Ser substrate where no cutoff distance is assigned to end the simulation. The distance being monitored is the distance between the alpha carbons of the X and prolyl residues. The force for this simulation is 60 kcal/mol*Å

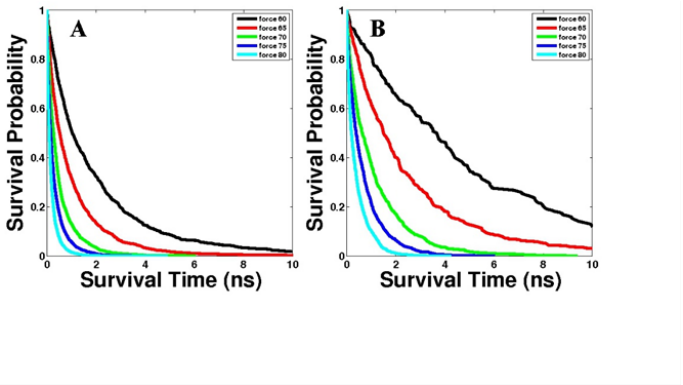

Figure 19. Survival Probability distributions for X = Ser (A) and X = pSer (B) substrates.

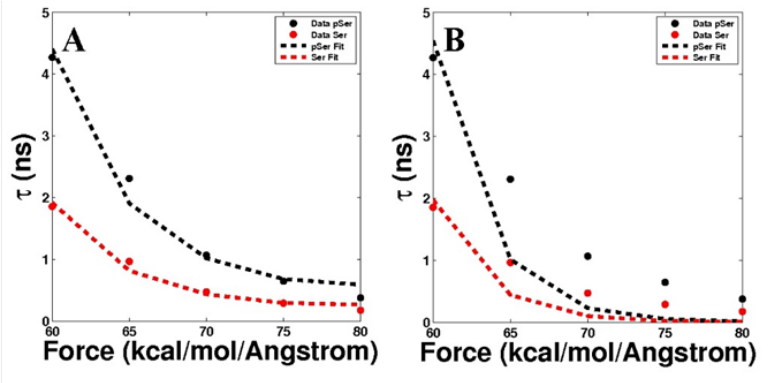

Figure 20. Kinetic fit of the constant force biased simulations to the DHS model (A) and the Bell model (B)

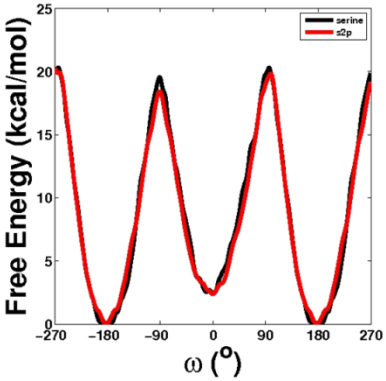

Figure 21. Free Energy Profiles of prolyl cis-trans isomerization for the X-Pro motif where X=Ser (black) or pSer (red).

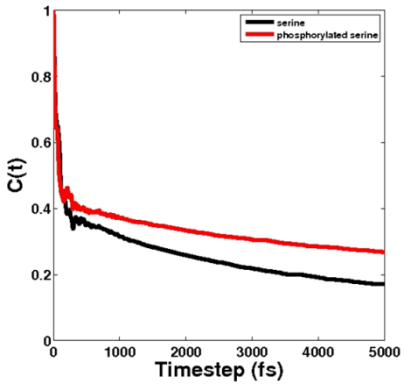

Figure 22. Autocorrelation function of unphosphorylated (black) and phosphorylated (red) polypeptide substrates at 300K.

4.8 Tables

Table 1. Data of constant pulling force simulations

X Residue	Force (kcal/mol*Å)	Number of isomerization events	τ (ns)
Ser	60	719	1.85
	65	1415	0.96
	70	2840	0.47
	75	4768	0.29
	80	7332	0.17
pSer	60	221	4.27
	65	434	2.31
	70	895	1.06
	75	1253	0.64
	80	1042	0.37

Table 2. Fitted Parameters for the DHS model

X residue	τ0 (s)	X‡ (Å)	ΔG‡ (kcal/mol)	υ
pSer	311.184	0.39	17.7	0.49
Ser	128.701	0.38	17.6	0.51

5 CONCLUDING REMARKS AND FUTURE STUDIES

The studies presented in this dissertation provide an atomistic level understanding of how prolyl cis-trans isomerization can occur both in vitro and in vivo to regulate subcellular processes. Pin1, a specific peptidyl prolyl isomerase, is used as a test system to understand how peptidyl prolyl isomerases circumvent the timescale issue in prolyl cis-trans isomerization to make it a tractable molecular switch. Pin1 does this through conformational selection of its substrates. A direct result of conformational selection for substrate selection is conformation directed catalysis. This method of catalysis yields noncovalent stabilization of the transition state through hydrogen bonding switches which create intermolecular interactions with the substrate and intramolecular interactions with Pin1 itself. In addition to transition state stabilization, these hydrogen bonding switches aid in binding of the ground states (cis and trans configurations of the ω bond) and maintain the structural integrity of the binding pocket. This work also serves to provide the scientific community with an elevated understanding of how enzymes can perform their catalytic function in a completely noncovalent manner.

Future studies involving Pin1 should center on furthering the knowledge of certain binding pocket residues. Currently, there is no experimental data of Cys113's pKa creating doubt about its protonation state over the course of a catalytic event. MM/PBSA calculations can be performed of the Pin1-substrate complexes to evaluate how binding differs when Cys113 is protonated and unprotonated. This study can also involve His59 and His117 which make up the dual histidine motif. Establishing protonation states for these residues can further the understanding how Pin1's hydrogen binding switches behave in the presence and absence of a substrate.

The studies presented herein also show how a mechanical force can be used as a substitute for a PPIase. It is believed that “mechanically” catalyzed prolyl cis-trans isomerization is how RNAse A and thioredoxin fold without a PPIase on the millisecond timescale. This is rather remarkable because a PPIase such as Pin1 is typically required to achieve a prolyl isomerization in the millisecond time regime. This particular aspect of the work presented has potential to applied in future studies for solving the protein folding problem. Protein folding is currently one of the grand challenges in science and holds the potential for revolutionizing medicine. If the protein folding problem can be solved, designing synthetic proteins to perform certain functions in the body becomes feasible. This creates a radical path for the development of cures and therapies for diseases.